We are Pfizer La Jolla

We Discover and Develop Innovative Medicines for a Healthier World

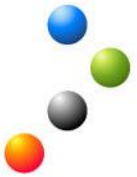
Leads Identification: Where Science Meets Art

Mehran Jalaie (MJ)

“All procedures performed on these animals were in accordance with regulations and established guidelines and were reviewed and approved by Pfizer Institutional Animal Care and Use Committee.”

Pfizer Inc.
La Jolla Laboratories
10614 Science Center Dr.
San Diego CA 92121
858-526-4065

Leads Identification



1. High Throughput Screen (HTS)
2. Parachute
3. Virtual Screening (VS) and Hit Follow-ups
    - Classical VS (high costs)
    - Structure-Guided Pharmacophore Method
    - Hit-Follow-ups
    - Current approaches (better and faster)

Most Projects Don't get HTS!

VS hits identified 20/55

VS applied 27/55

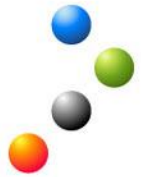
Concept	ESD			SDS			LD	CS		BU
	1	2	3	4	5	6	7	8	9	10
CNS		mGluR4	STEPi	mGluR4 PAM	MT-1/2		5HT6 antag	Alpha2a		NRI
				GABA B Mod	GABA-A2		D2/5HT2/SRI			D2 PA
				a2d subtypes	Nav_lowD2		D2PA/SRI			
					Lamotrigine+		NRI/5HT1a			
AB				HPPK	BC		NRI/D4			
				MurG	PheRS		LpxC			
				GlmU			GyrATPase			
					Smo Ag					
DERM		GSK3		MC5R	S1P		Alk5			AR
				PDGF	SCD-1					
							NMDA			
INF			BTK	PBR	PAD4		IL6-mAb			
			C1S	cPLA2a	PDE2		FAAH			
				IRAK4	S1P1					
CV		FLAP	PCSK9	F11a	SKY		LTCC			ARB+
			PAR1	F9a	SPT					
			SQS-i	P2Y1	NPC1L1					

27 programs have applied virtual screening

20 programs have identified chemical matter by virtual screening

No HTS?

Classical Virtual Screening



- Virtual Screening: what it used to be …
  - PfiSearch / other 2D methods
  - ROCS/EON or Pharmacophore

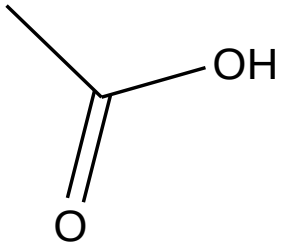
- Success stories at high costs:

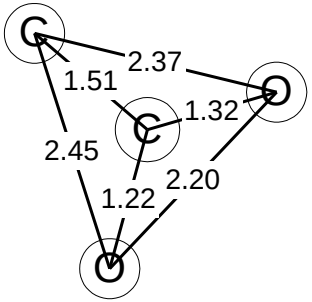
– Lamotragine+	no chemistry	CAN (Candidate Alert Nomination)
– LTCC	no scaffold	CAN

PfiSearch Similarity Search

3D Molecular Graph

•Vertices are defined as atom types
•Edges are defined as distances or distance ranges



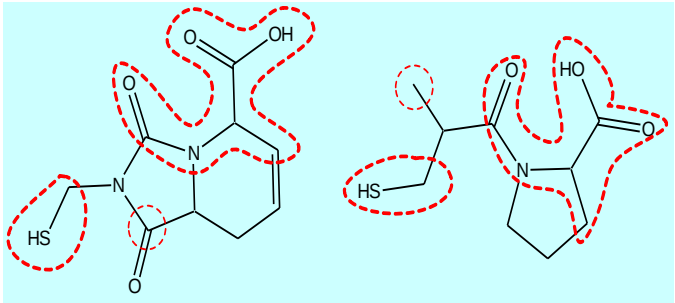


•Distance Geometry

Establish the graph edge distance bounds as the minimum and maximum allowed distances between two atoms

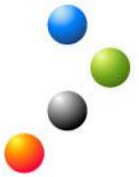
3D MCS

•Largest subset of atoms common to both Structures such That:
•Atom types are compatible
•Atom-pair distance ranges in both structures are compatible (*i.e.*, overlap to an acceptable degree)



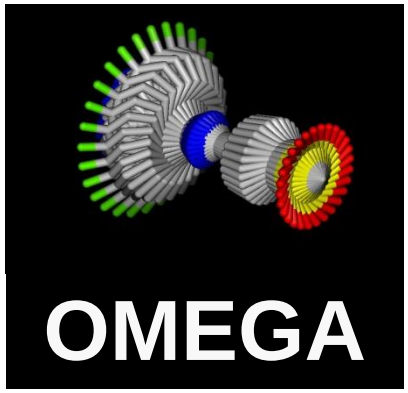
$$S_{12} = \frac{|G_{12}|}{|G_1| + |G_2| - |G_{12}|} = \frac{11}{16 + 14 - 11} = 0.58$$

ROCS



- ROCS (**R**apid **O**verlay of **C**hemical **S**tructures)
- Rigid shape-based superposition
- The molecules are aligned by maximizing the overlap volume
- Not using hard sphere representation
- Can use overlap of functional groups (donors, hydrophobic…): Pharmacophore on the fly.

The “Shape” of Ligand-Based Design



OMEGA

pre-generated rigid multi-conformations

Probe	
Co-Crystal	Ligand-Based
Docking	Hypothesis

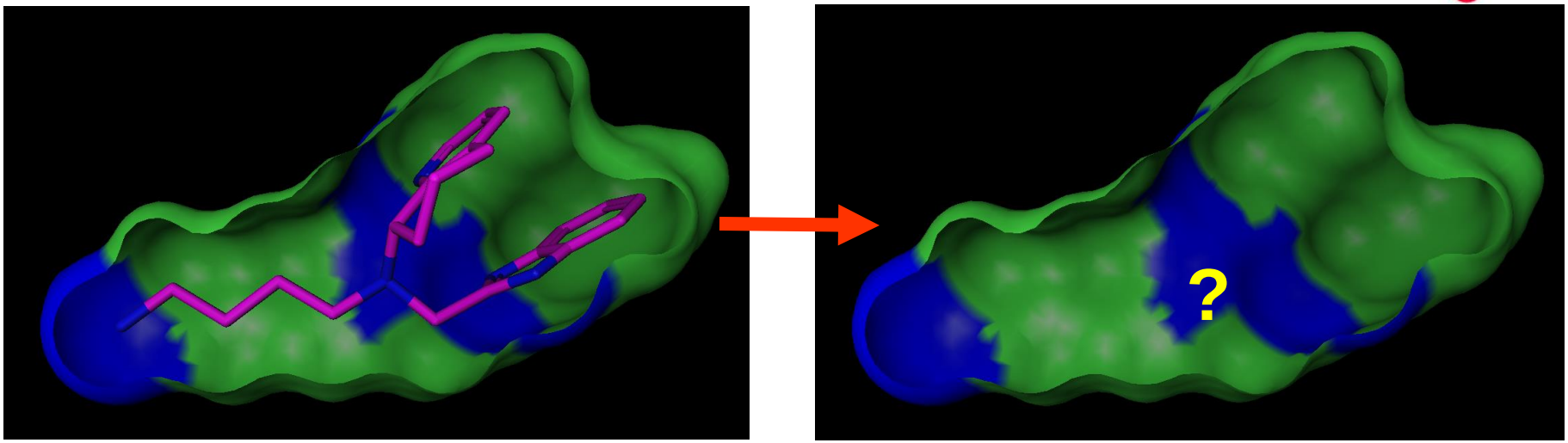
ROCS

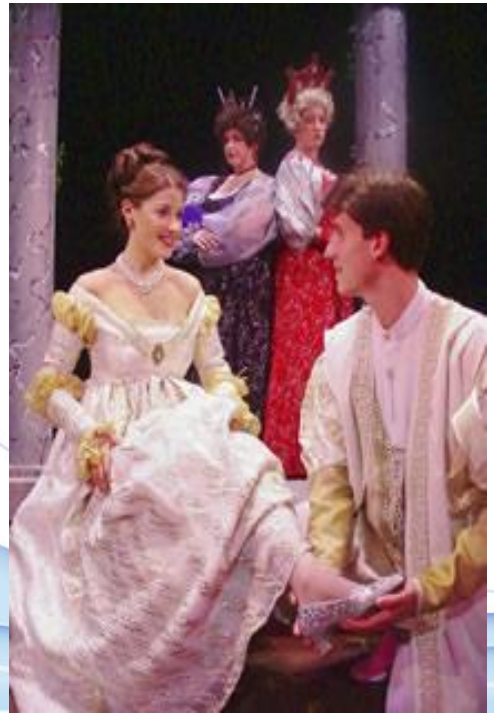
shape similarity

EON

electrostatic similarity

Cinderella's Slipper





"The Prince picked up her slipper
and said to his ministers, "**Go and search everywhere for the girl whose foot this slipper fits**.
I will never be content until I find her!"
So the ministers tried the slipper on
the foot of all the girls... and on Cinderella's foot as well...
Surprise!
The slipper fitted perfectly…."

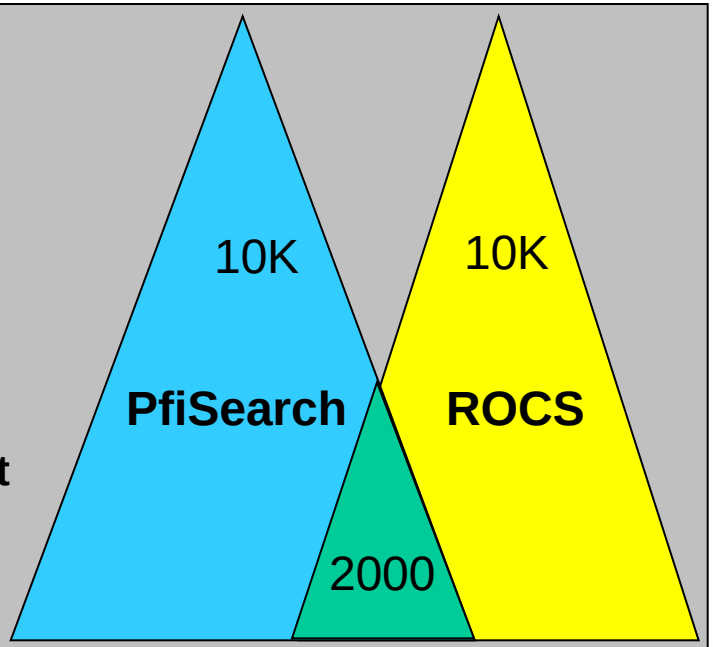
From the Fairy Tale of Cinderella

Classical Virtual Screening Tools

- Ligand-based Approaches:
  - PfiSearch
  - ROCS/EON(3D)

How it works

- **ROCS/EON**
- **PfiSearch / 2D searches**
- **The overlap between both methods**
- **Each method has its unique strength and merit**
- **The top ranking hits from each method should be tested as well**

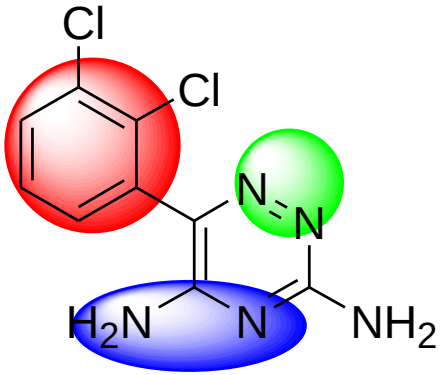


The First Law of Virtual Screening

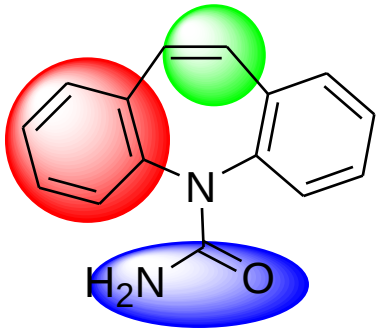
Genuine active hits **must exist** in the screening collection before you can identify them

Lamotrigine+ project

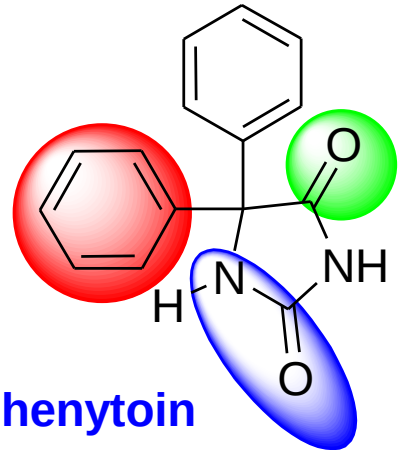
- For the project, a source of novel chemical matter required for several reasons: a) >10 companies pursuing similar product profile (personal communication), b) lamotrigine chemotype well explored by GSK and others, therefore lamotrigine-like chemo-space will be/get crowded
- In addition to using lamotrigine (control seizures) as a 3D probe, carbamazepine and phenytoin (bipolar disorder) would provide greater diversity in the identification of chemical matter.



lamotrigine



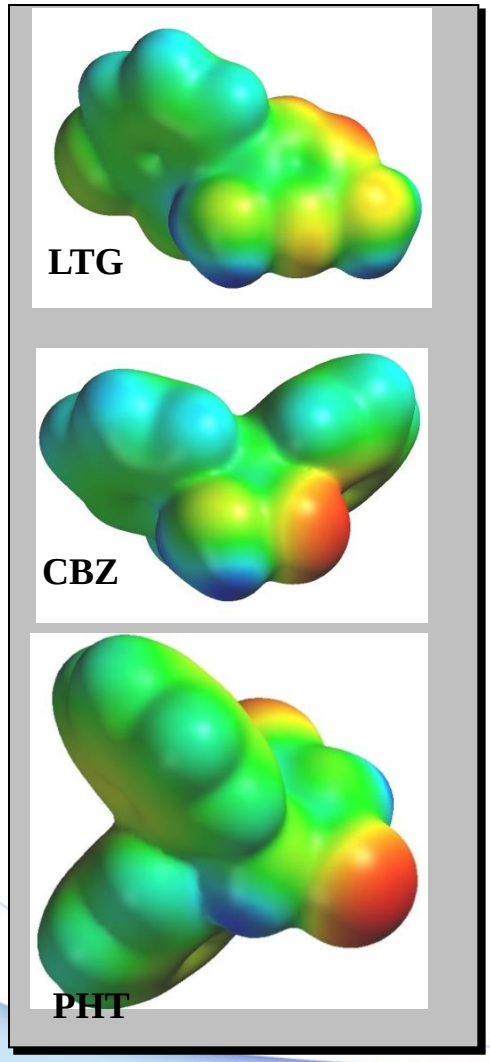
carbamazepine



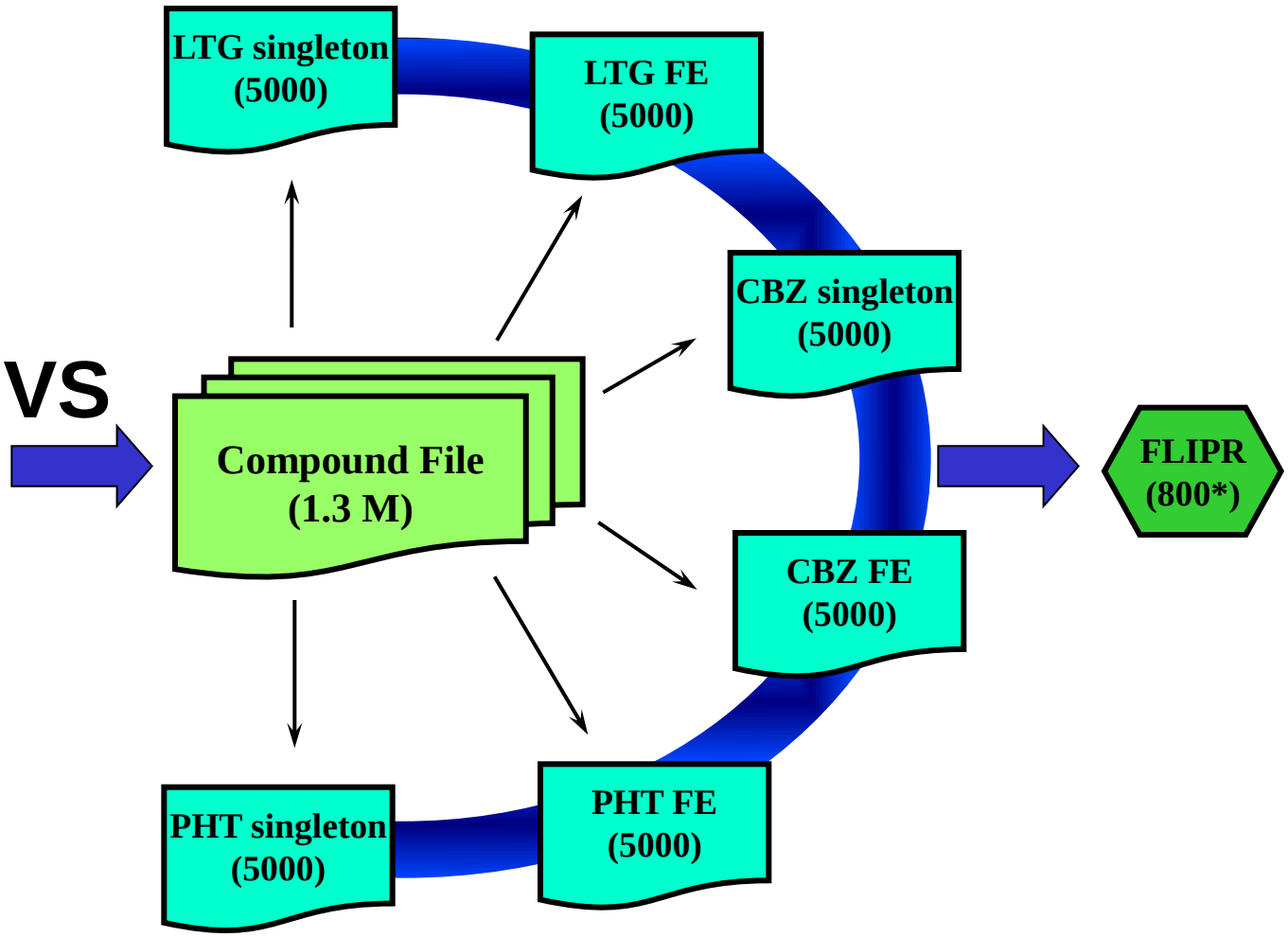
phenytoin

The Project Achieved Its Goal: High Costs

800/30000 = 2.6 %

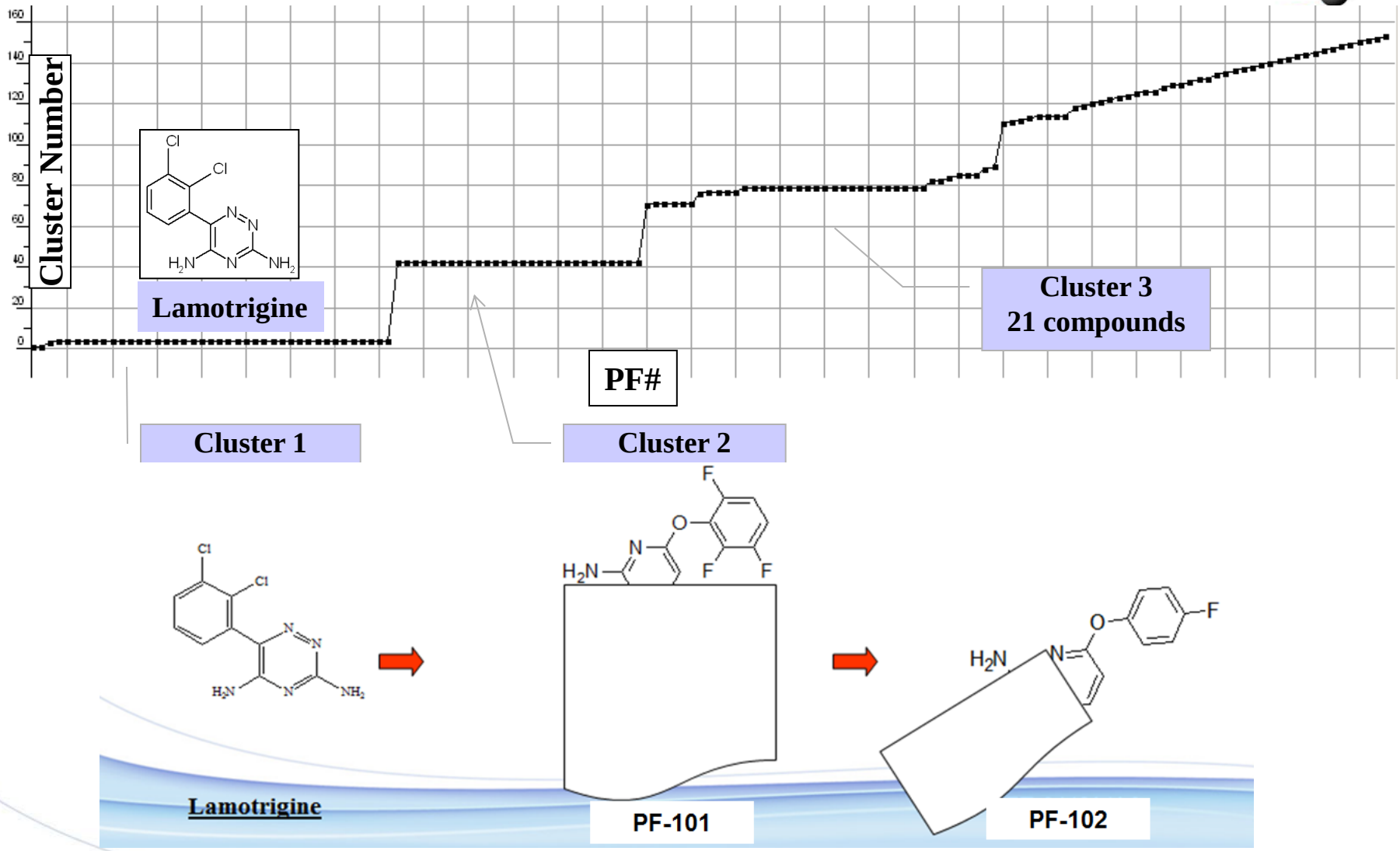


VS



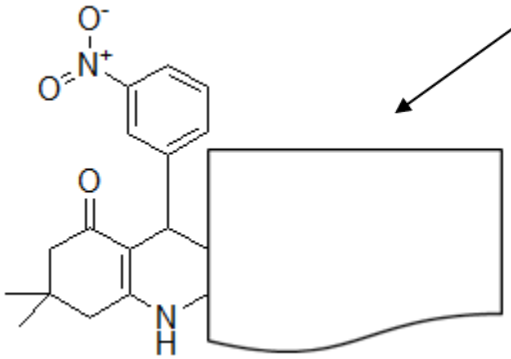
*Hit = % inhibition ≥ 50 at 30 and 10 μM

153 Confirmed Hits From ~ 5000 VS

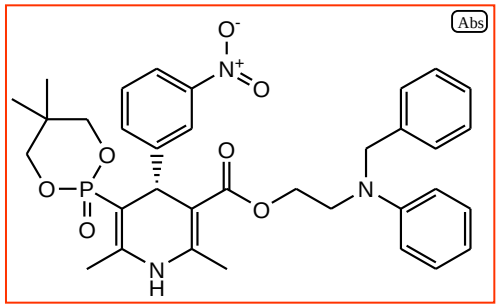


L & T Calcium Channel Blocker (LTCC)- Hypertension: PF-103 and PF-105

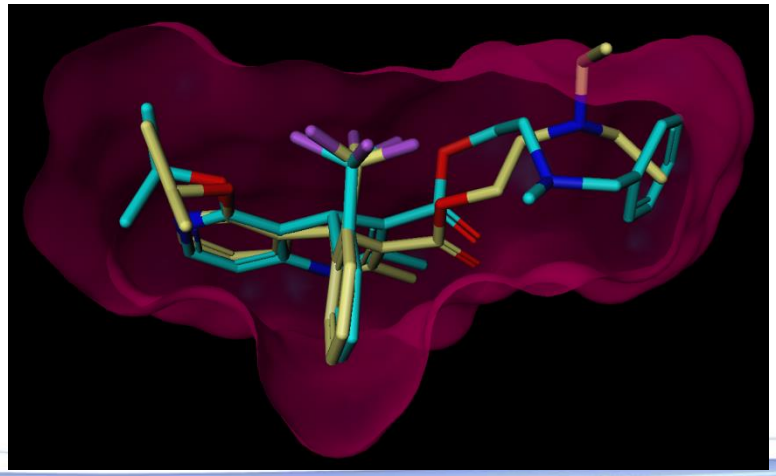
1. Virtual screen was initiated with known L-type calcium channel antagonists probes.
2. Looking for different scaffolds with better PK with L and T activities:
3. 5000 compounds were tested
4. 150 compounds were found to have confirmed activity (150/5000 = 3.0 %)

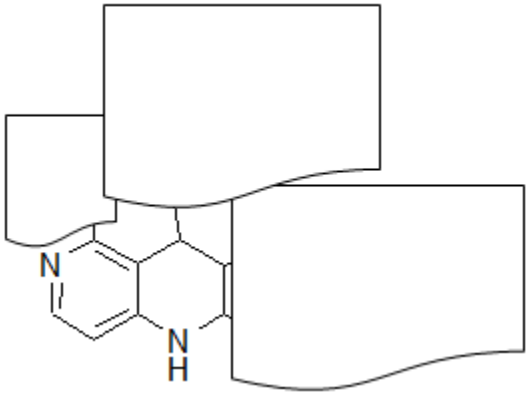


PF-103
PAA = 3000 nM
T-FLIPR = 71 nM
L-FLIPR = 378 nM



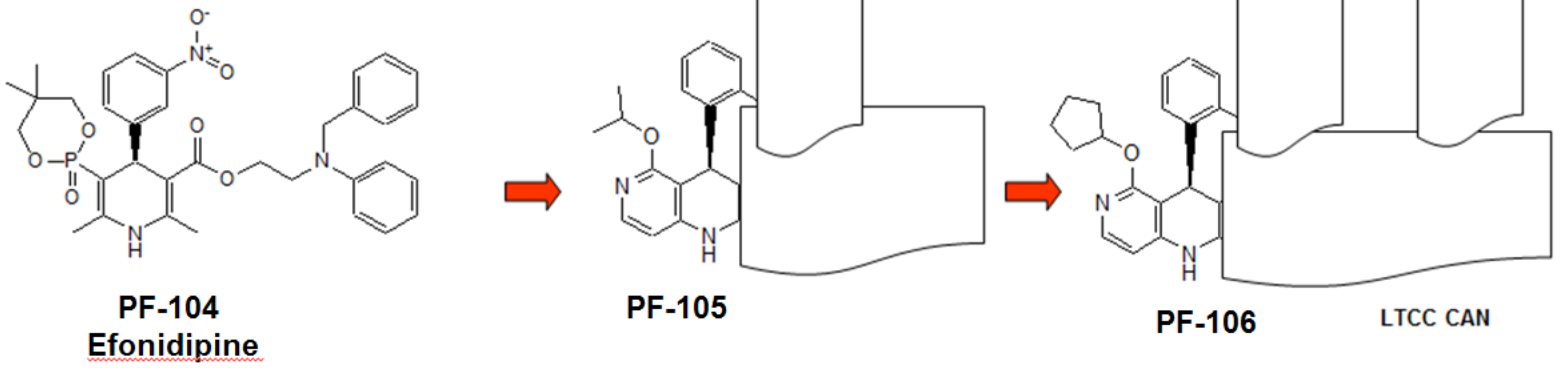
PF-104
Efonidipine





PF-105
PAA = 4 nM
T-FLIPR = 557 nM
L-FLIPR = 220 HLM = 15 min
Efficacy in SHR: @ 30 mpk PO lowered BP ~75mmHg without changing HR. Duration of efficacy (> 12 h) !

PF-106: Chemistry focused on the DHN series



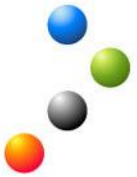
PF-104
Efonidipine

PF-105

PF-106 LTCC CAN

Structure-Guided Pharmacophore Method

PKC (Protein kinase C)



Therapeutic hypothesis:

Inhibition of PKC kinase has the potential to:

- Block tumor growth, invasion and metastasis

Specific Goals:

Develop small molecules capitalizing on the crystal structure and existing chemical equity:

- Potent, selective inhibitors of PKC with favorable overall properties for an oral drug
- Potent suppression of tumor growth and metastasis *in vivo*
- Little, if any effect on normal differentiated cells

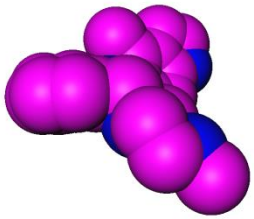
Goal

- Identify novel hits prior to HTS by testing less than 300 compounds

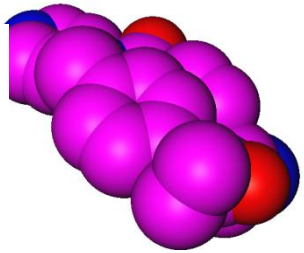
Considerations

- Not interested in traditional Virtual Screening methods: how do we go about from @7 million compounds to 300 compounds with potential lead qualities for PKC?
- Using existing Co-x'tal effectively to select compounds

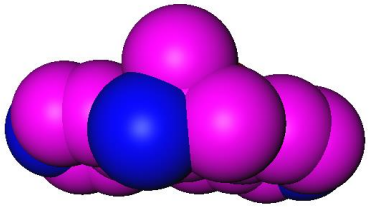
Why a Structure-Guided Directional Pharmacophore?



PF-107
Ki = 0.98 nM
clogP = 4.8
MW = 464

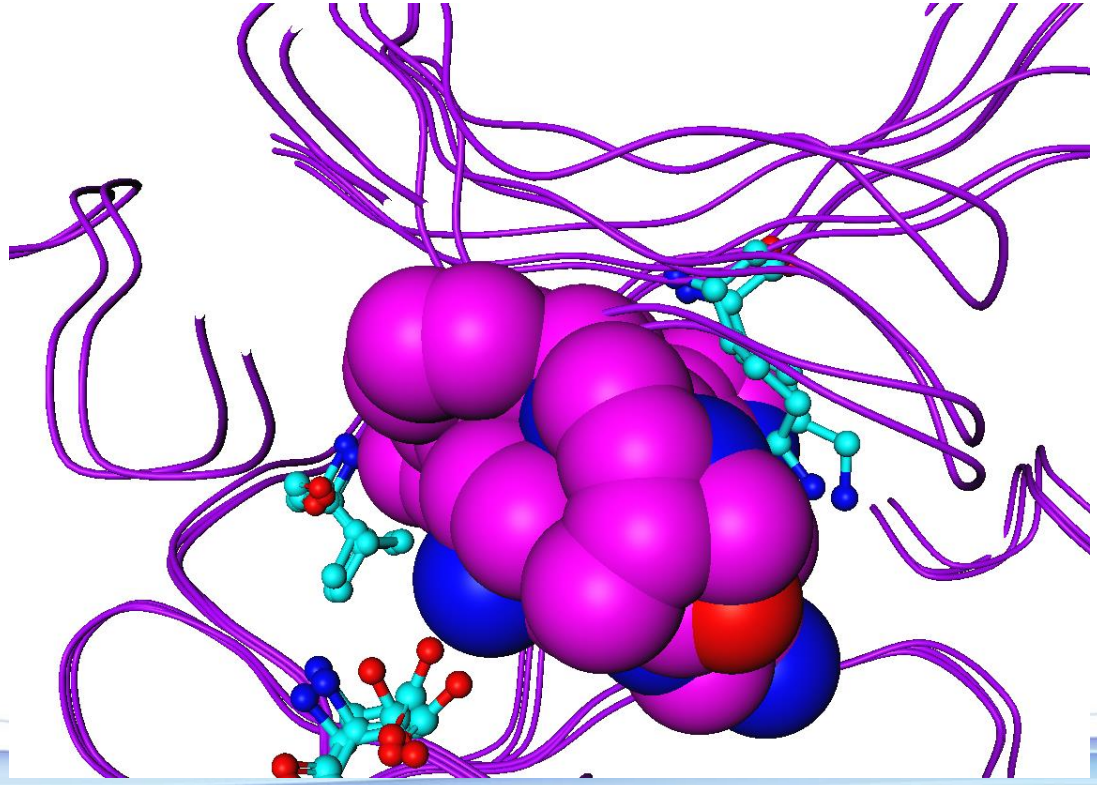


PF-108
Ki = 0.18 nM
clogP = 2.8
MW = 434

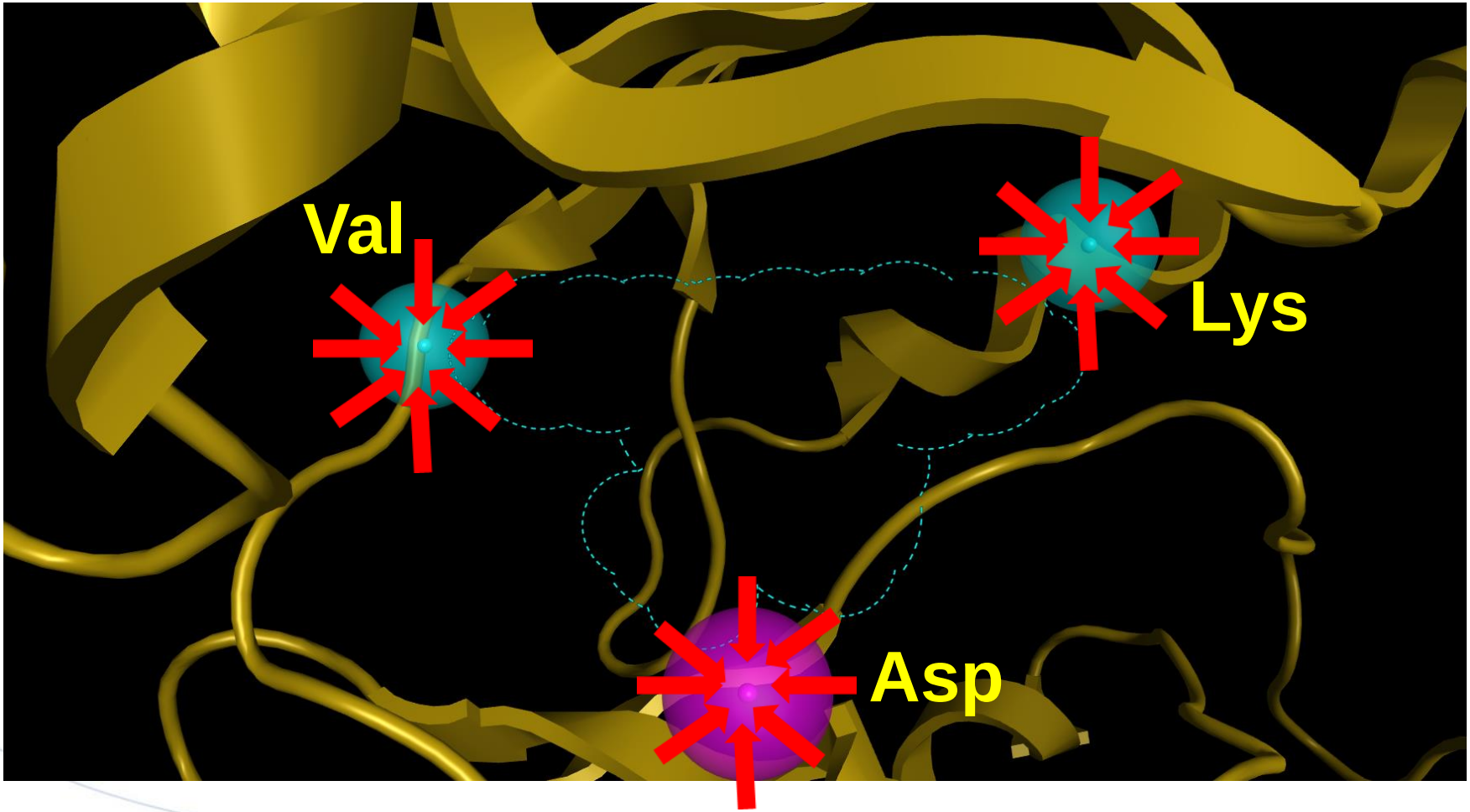


PF-109
Ki = 16.6 nM
clogP = 2.2
MW = 297

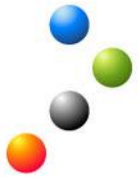
The overlay of all the co-crystal structures



The Directional Feature Pharmacophore & Excluded Volume



The Strategies

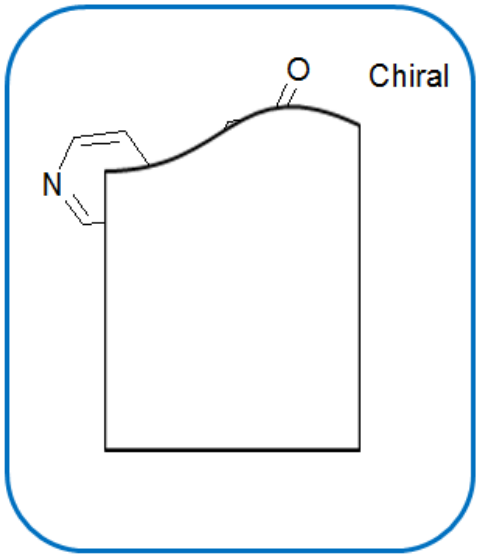


- Built a structure-guided pharmacophore using **directional** features
- Searched Pfizer collections of 7 million compounds with the structure-guided pharmacophore
- **5628** matched 3 features and the excluded volume
- **5628** docked into 3 conformations of the enzyme without water molecules
  - One feature must be maintained to the hinge
  - Filtered heavily (MW< 450, cLogP 3.5, acceptable Strain Energy and Availability): **308** Hits.
- **308** triaged to **235** sent for testing

Results: 58 active / 235 tested = 24.7%
44/235 = 18.7% (Ki < 1uM)

PF-110

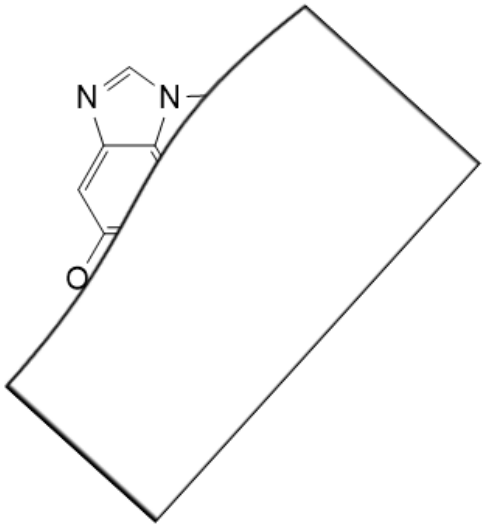
Ki =300 nM



109 co-crystal

PF-111

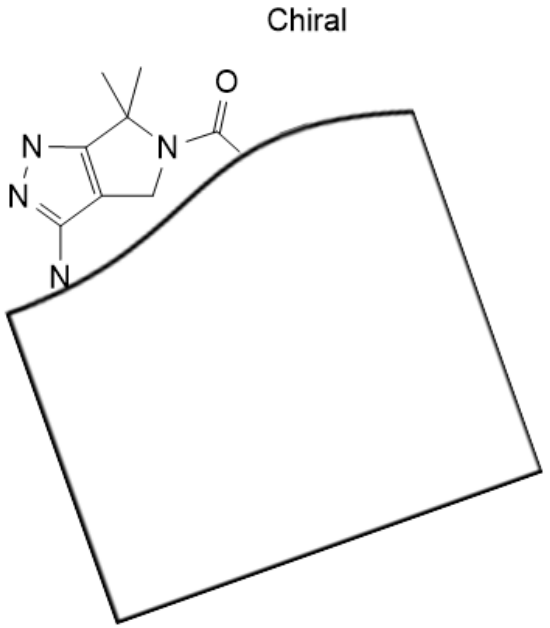
Ki = 75 nM



1 co-crystals available

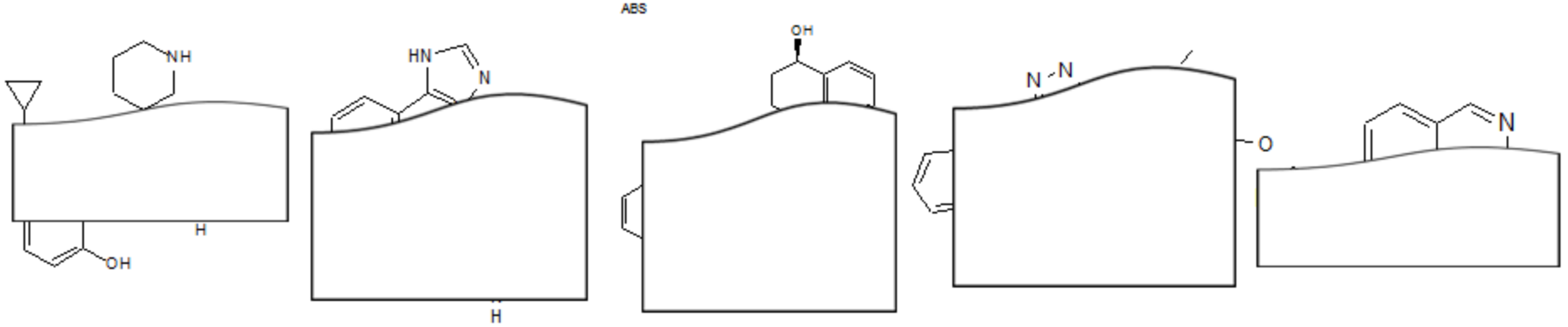
PF-112

Ki = 11 nM



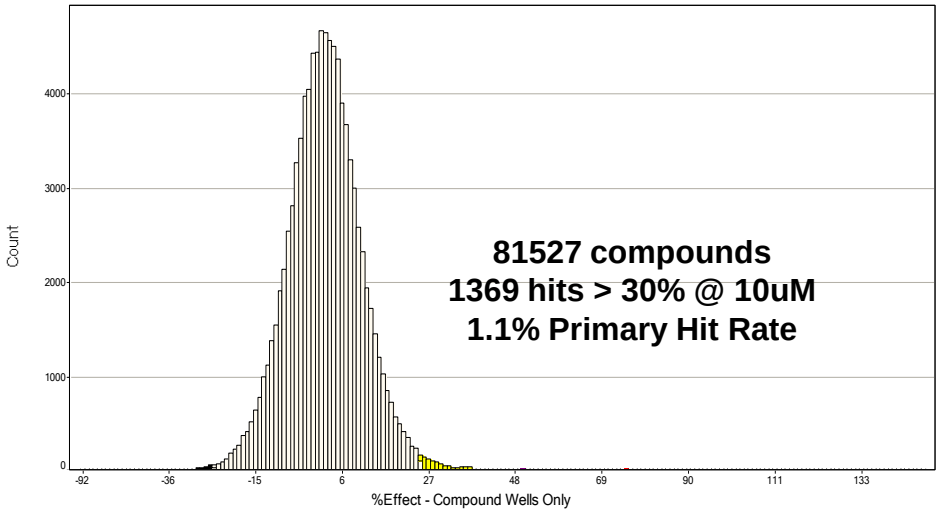
26 co-crystals available

Emerging Pharmacophore Hit Series II



PF-113	PF-114	PF-115	PF-116	PF-117
Ki =18 nM	Ki = 135 nM	Ki = 30 nM	Ki = 594 nM	Ki = 3.5 nM
The exact co-crystal available	1 co-crystals available	7 co-crystals available	21 co-crystals available	11 co-crystals available

HTS: 1017/66226 = 1.1 %

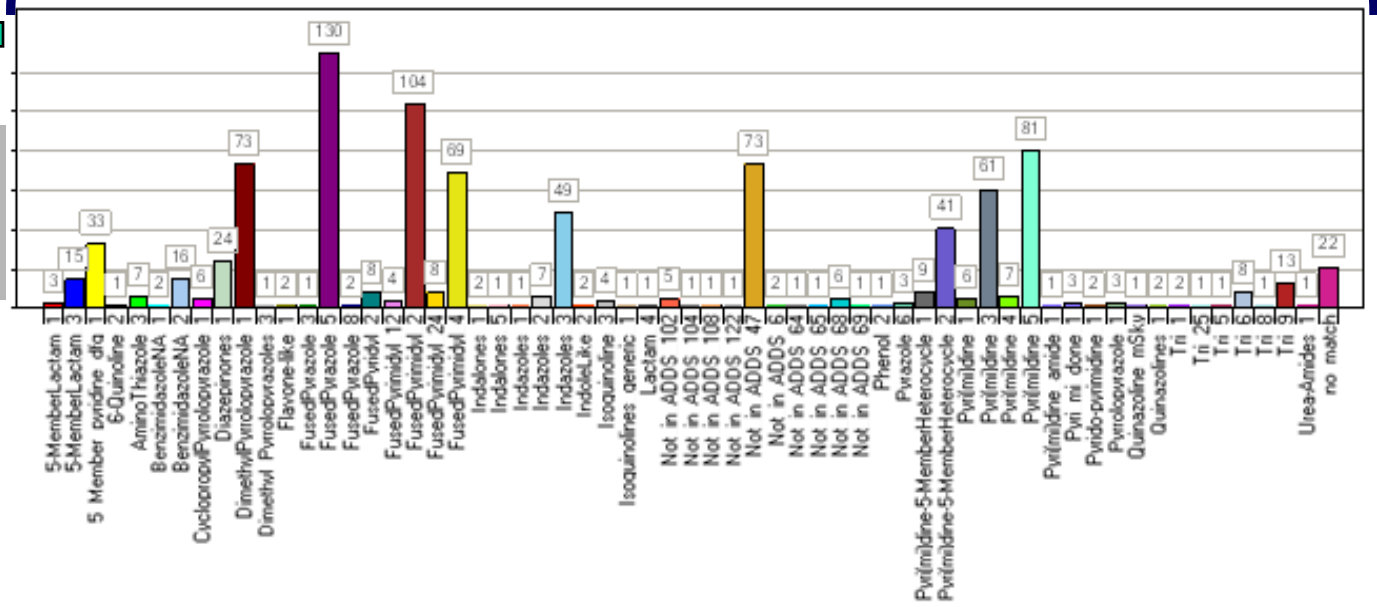


Collection	~# of Compounds	# Confirmed Hits	Confirmed Hit Rate (%)
KTL 2.0	36409	500	1.4
KTL 1.2 Classical	15633	411	2.6
KTL 1.2 Non-Classical	12918	14	0.1
GDRS2 Subset	12664	32	0.3

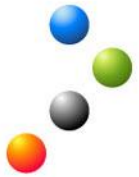
Overall = 1.1 %

Chemistry team triage eliminated >50% of confirmed hits

Collection	# Selected for KGP IC50	%
KTL 2.0	319	64
KTL 1.2 Classical	83	20
KTL 1.2 Non-Classical	4	29
GDRS2 Subset	18	56



Summary



Pharmacophore

- Series I & II along with several other class of compounds were Identified
- 235 compounds tested

Traditional Hit Exp

- Traditional hit expansions drove the activity to **33 nM** and identified new hinge binder motif
- 567 compounds tested

Assay-based Hit Exp

- Assay-based screening identified new scaffolds and hinge binder motif and drove the activity to **8 nM**
- 539 compounds tested 35 compounds < 1 uM.

HTS-Summary

- **Pharmacophore resulted in <u>24% hit rate</u> with and establishing the lead series.**
- **HTS <u>hit rate 1% and no found no additional series.</u>**

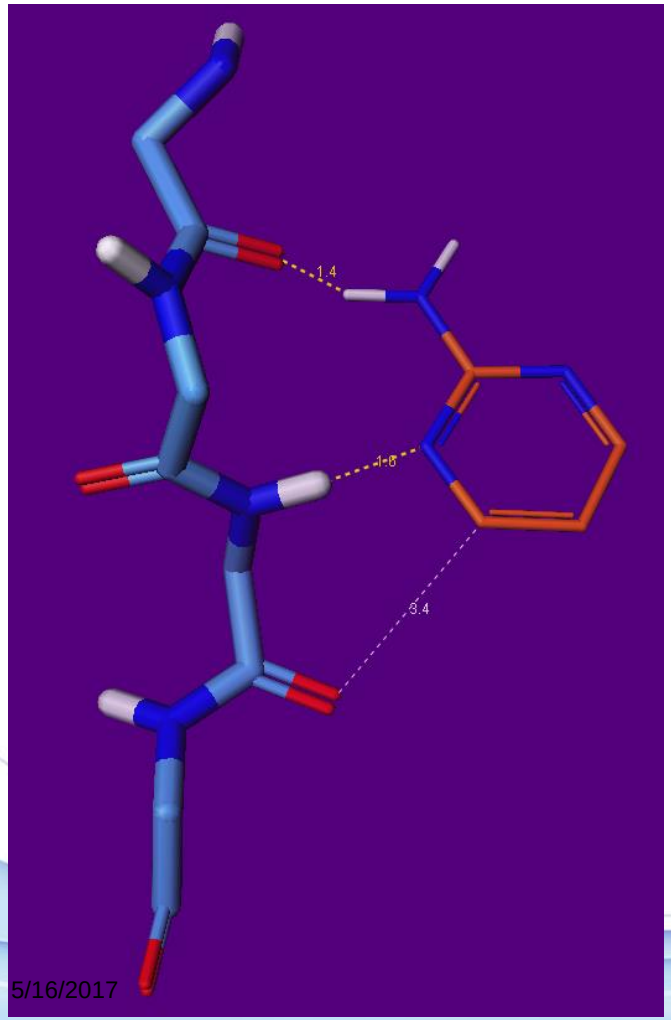

Hinge Swapping Strategy

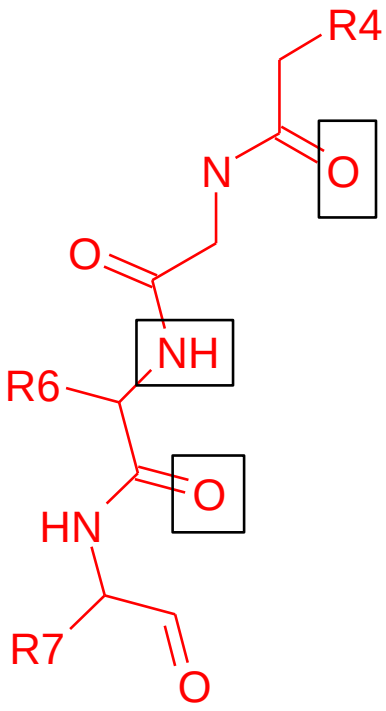
The Strategy

- Crystal Structure Database Base (CSDB) consists of thousands of kinase complexes.

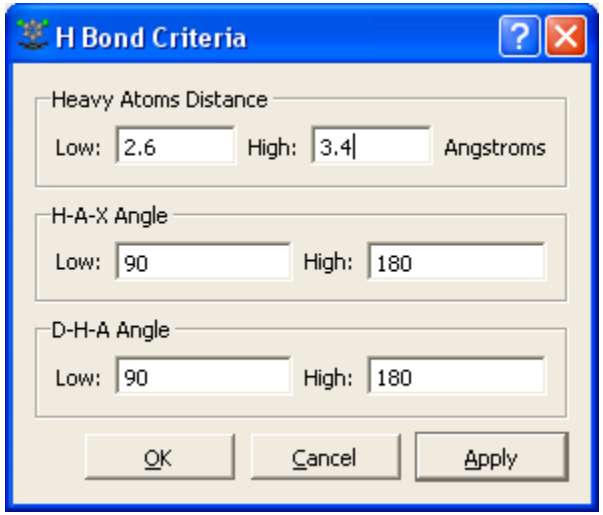
- We extracted the kinase complexes.
- Hydrogen bonds geometries were analyzed for identifying the “hinge binders”.

For Each Kinase Co-crystal Structure in CSDB, Determine Whether or Not it Forms at Least One H-bond with the One of the Hinge Residues(H4, H5, H6 and H7)



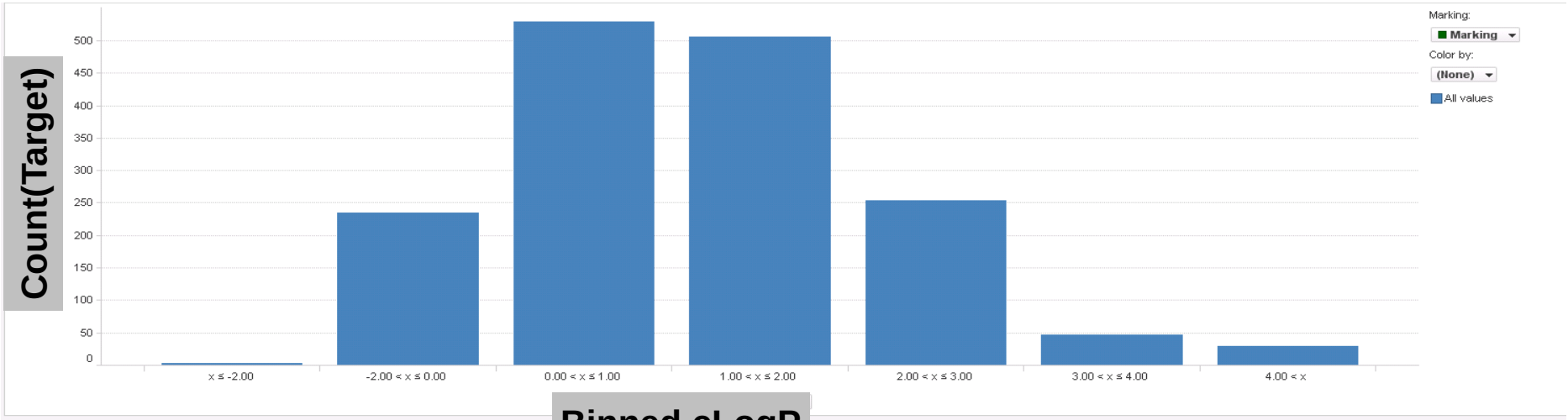


Hydrogen bonds between the ligand and the hinge residues are determined based on geometric criteria:



- Once a ligand was found to form H-bond to a hinge residue, the hinge binding fragment was extracted.
- If a ring fragment directly formed hydrogen bonds with the hinge, all ring atoms were extracted.
- In case the H-bond was through a non-ring atom, the algorithm found and extracted the nearest ring fragment in addition to the non-ring atoms that form the H-bond.

Output is a Collection of Unique Hinge Fragments

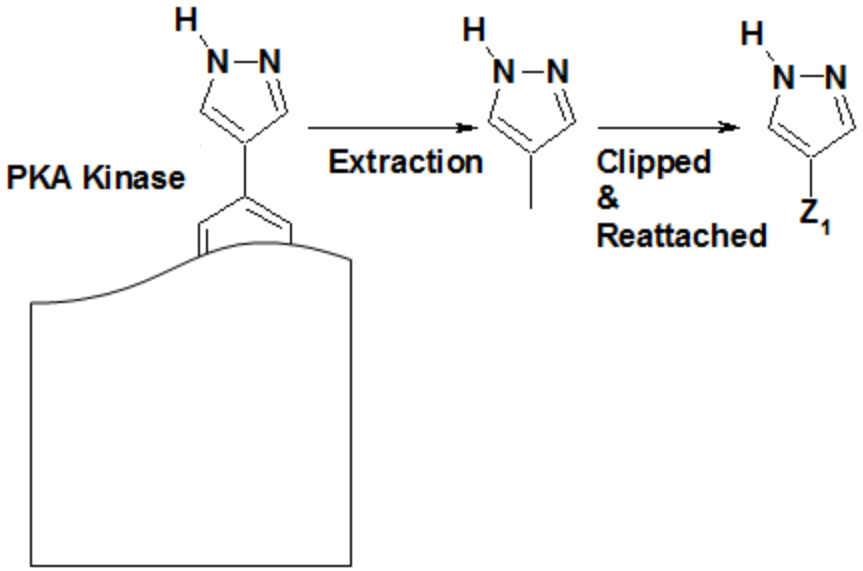


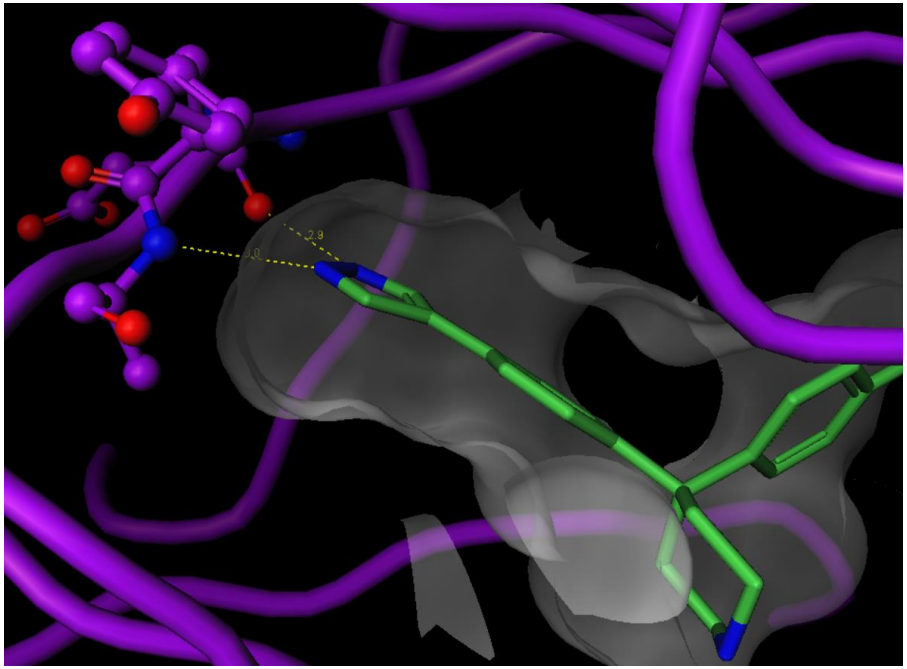
Binned cLogP

Targets

There Were 920 Unique Hinge-Binding Structures and the Clipping Points Were Preserved

- For example, the source of pyrazole as a hinge binder was 2UW7 (PKA kinase).

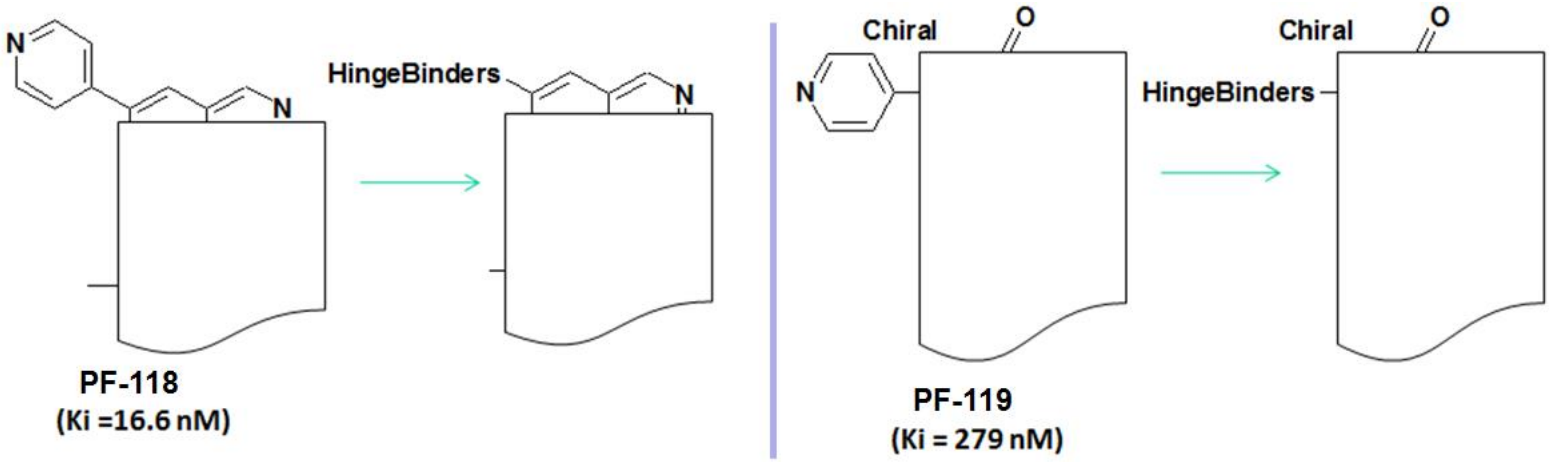




Re-attached the Extracted Hinge to a Core of Our Choice

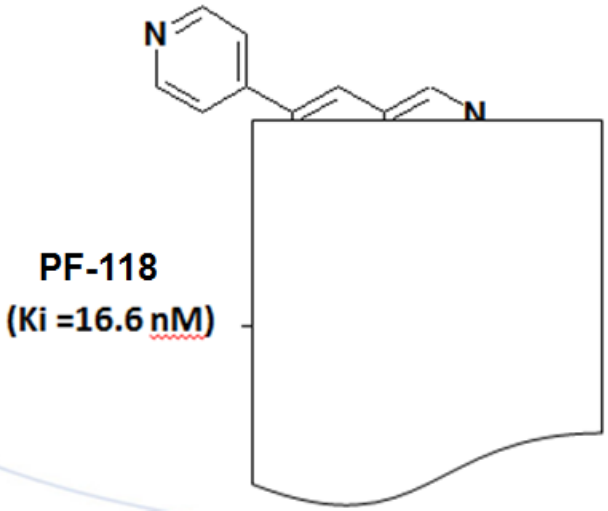
920 Hinge Binders were re-attached to two cores below from PF-118 and PF-119.

Applications of physical property filters (MW < 500 and cLogP < 3.00) reduced the number of enumerated compounds to 618.

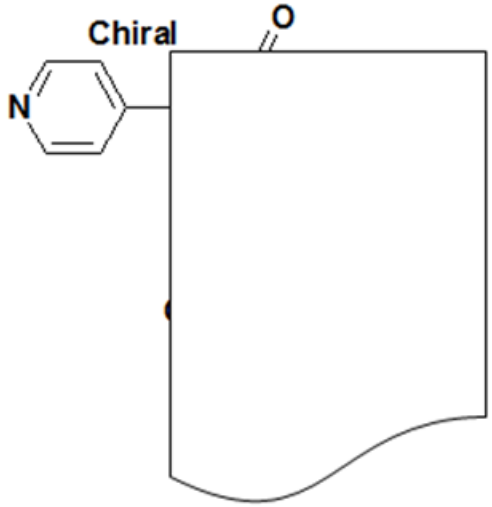


Modeling

- All the enumerated products were docked by Glide SP into the co-crystal structure of PF-118 in the absence of water molecules.
- All the docked poses were rescored with HT, and the local strain energies (AMBER and OPLS2005) were calculated.
- After removing strained* molecules and focusing on molecules with high calculated LE (< -0.2), 67 hinge binders remained.
- Based on the score, visual inspection and the level of the risks (selectivity), only a handful of hinges were selected for synthesis.



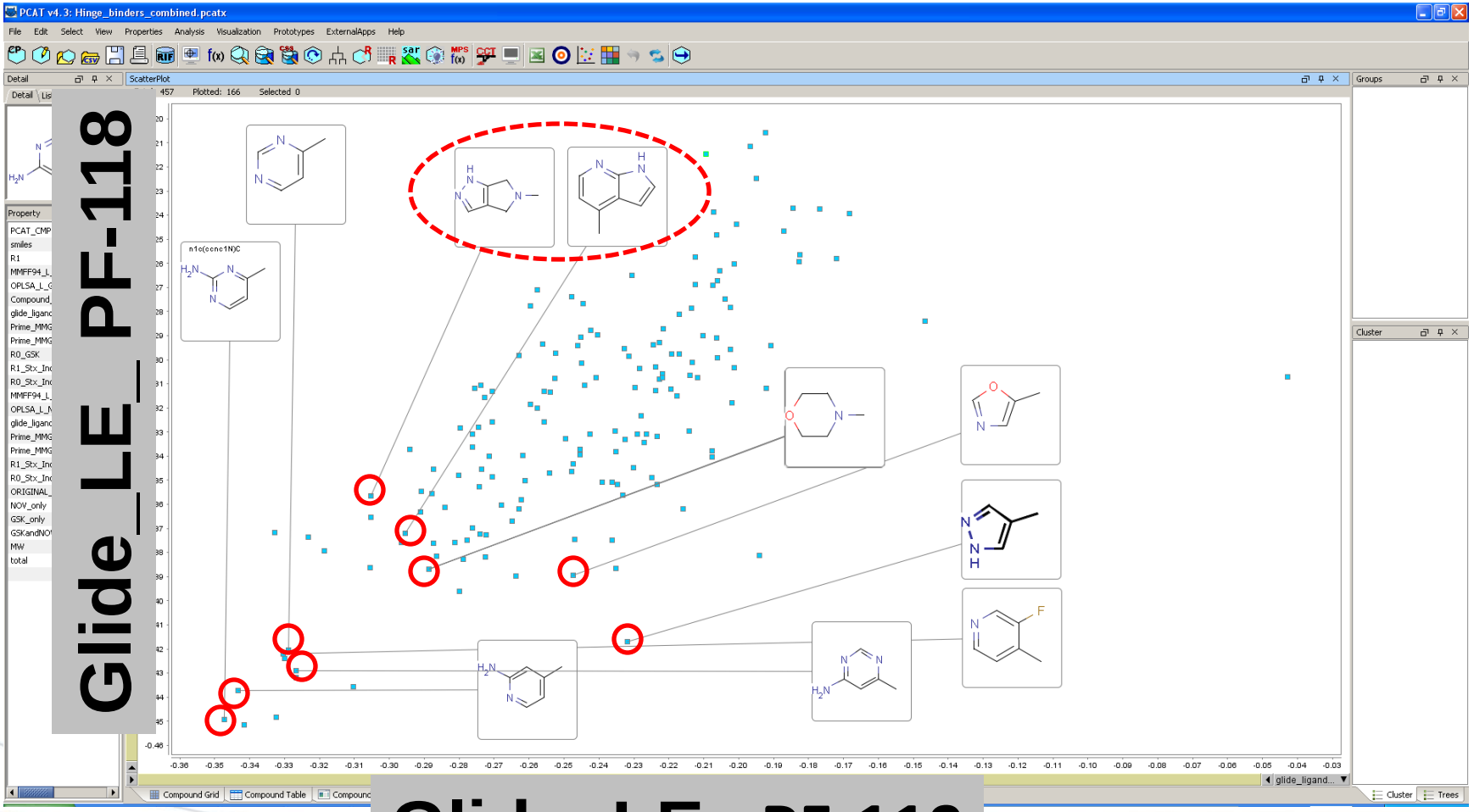
PF-118
(Ki =16.6 nM)



PF-119
(Ki =279 nM)

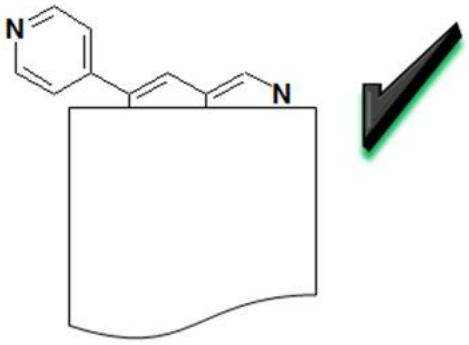
* Charifson et al; J Med Chem. 2004 May 6;47(10):2499-510

Interesting Hinge Binders Score/Visual

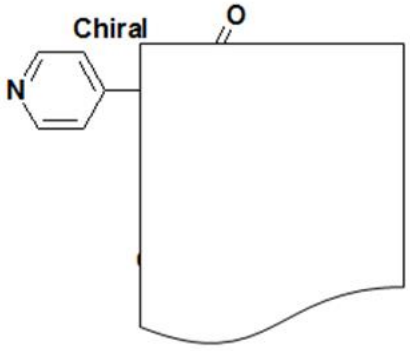


Reality …

- Series of singletons were followed up to test the idea of hinge replacement for the above cores.

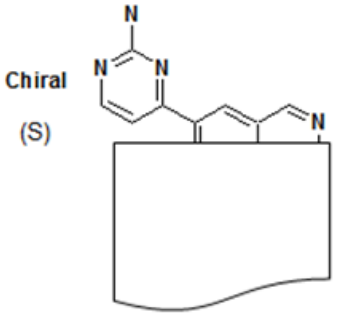


PF-118
(Ki =16.6 nM)
Series I

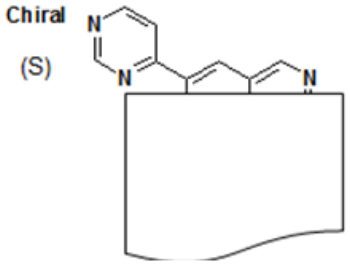


PF-119
(Ki =279 nM)
Series II

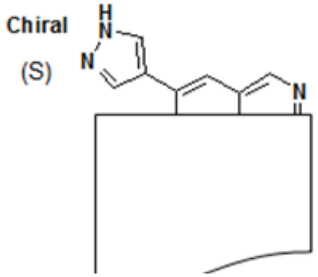
Results of Hinge Swap



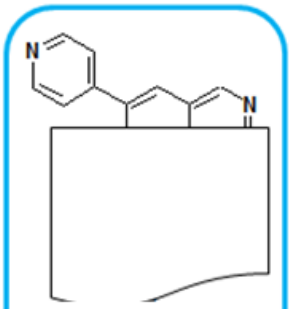
PF-120
1.05 nM



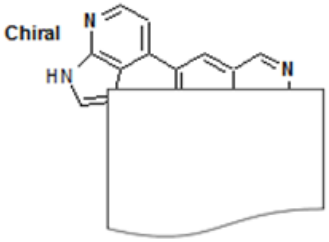
PF-121
1.17 nM



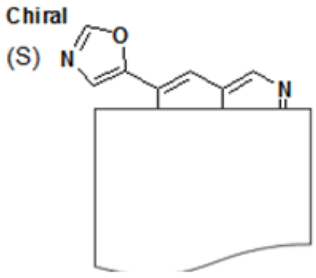
PF-122
16.4 nM



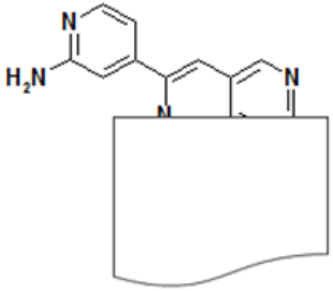
PF-118
16.6 nM



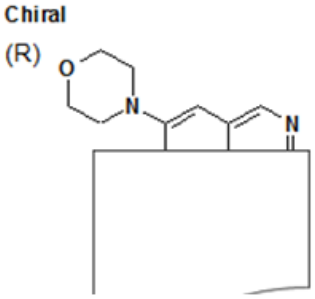
PF-123
0.25 nM



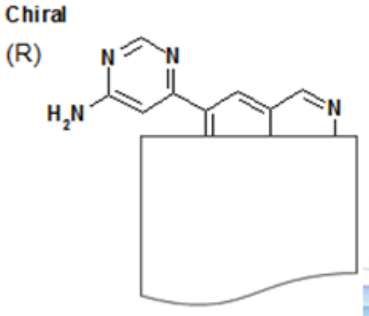
PF-124
20 nM



PF-125
85.9 nM



PF-126
93.0 nM



PF-127
295 nM

Conclusion

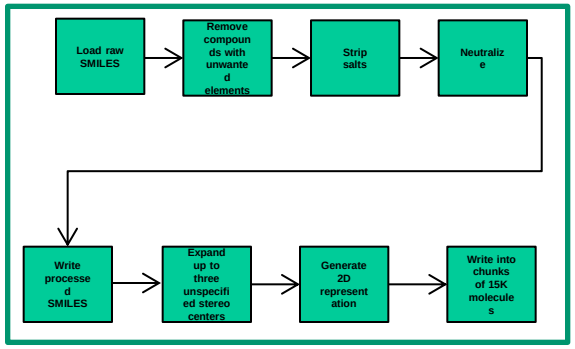
- Successfully identified more than 8 new hinge binders using a practical approach
- Extracted and re-attached existing hinge binders our lead series
- The new hinge binders are diverse and potent for our series.

Virtual Screening: Better & Faster

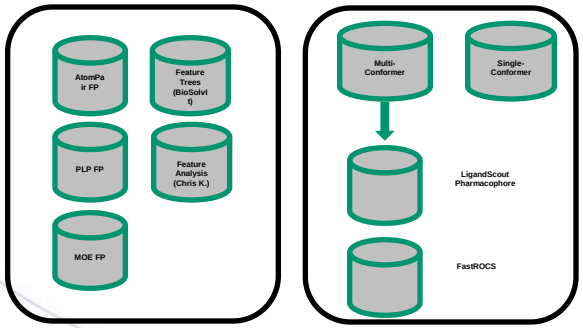
Virtual Screening Infrastructure

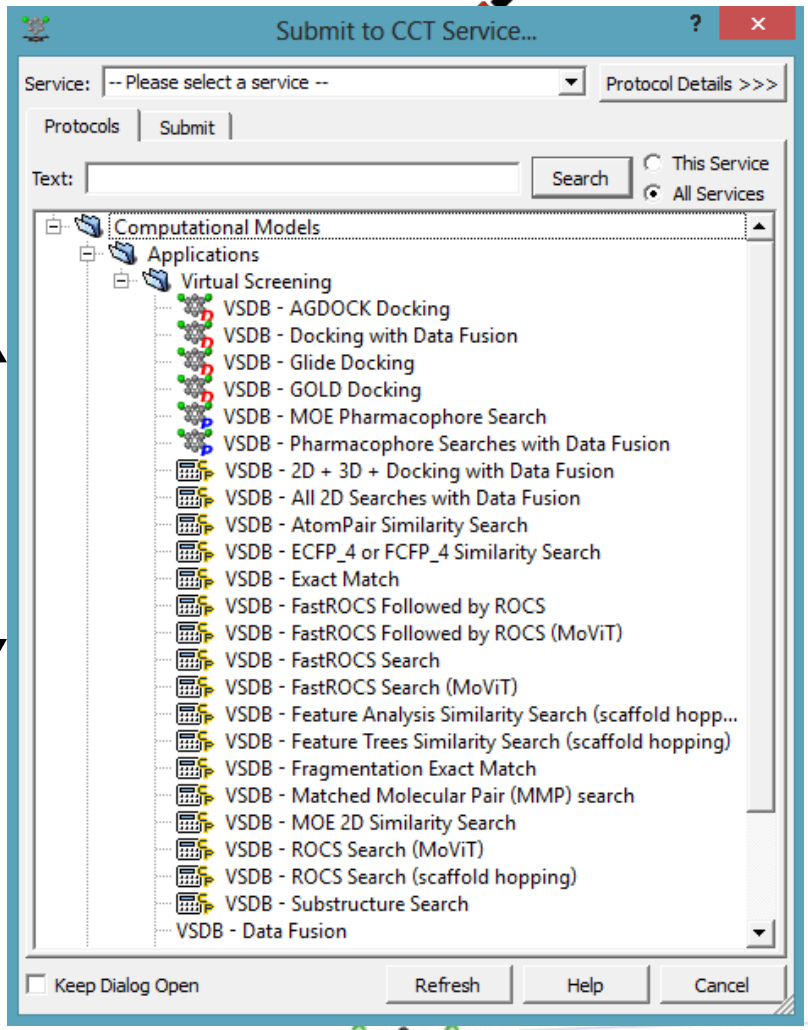
CCT

Carefully curated databases



Derived databases





Dockings

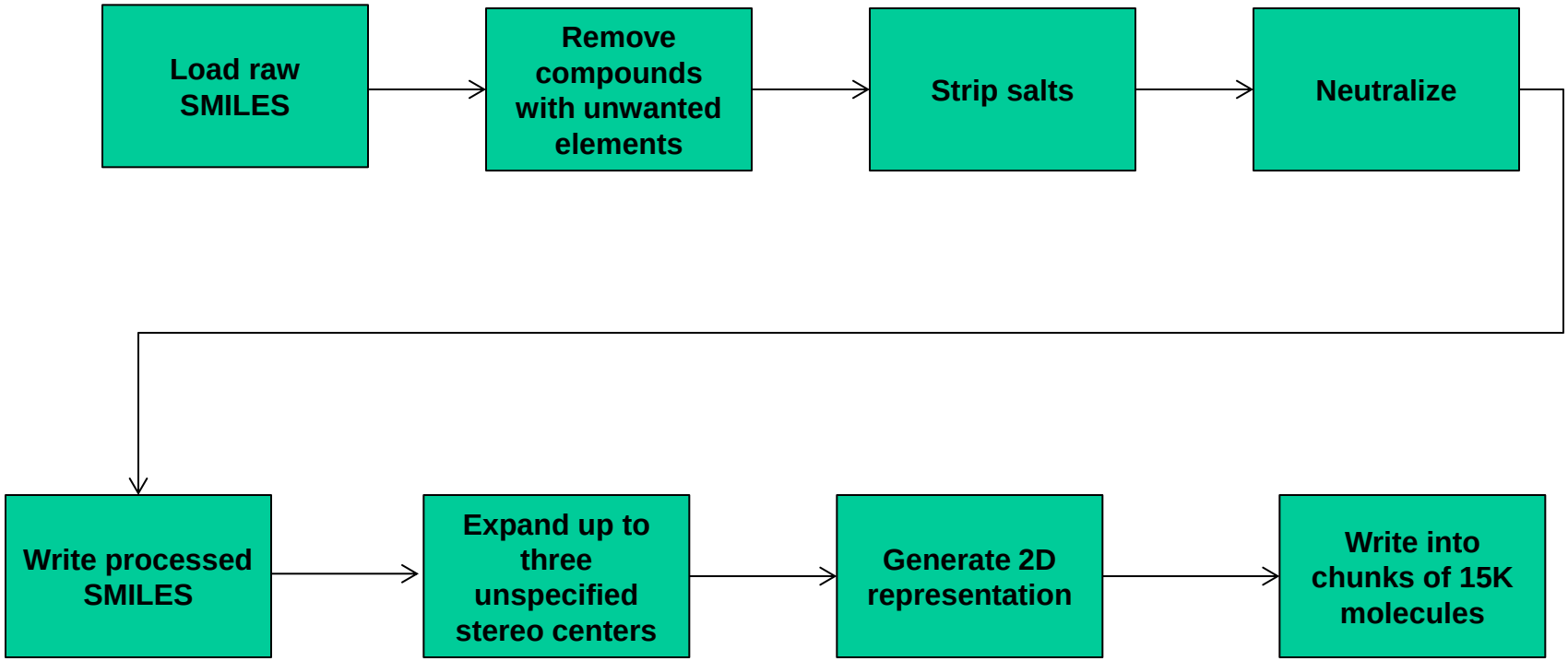
Pharmacophore

FastROCS/ROCS 2D Searches

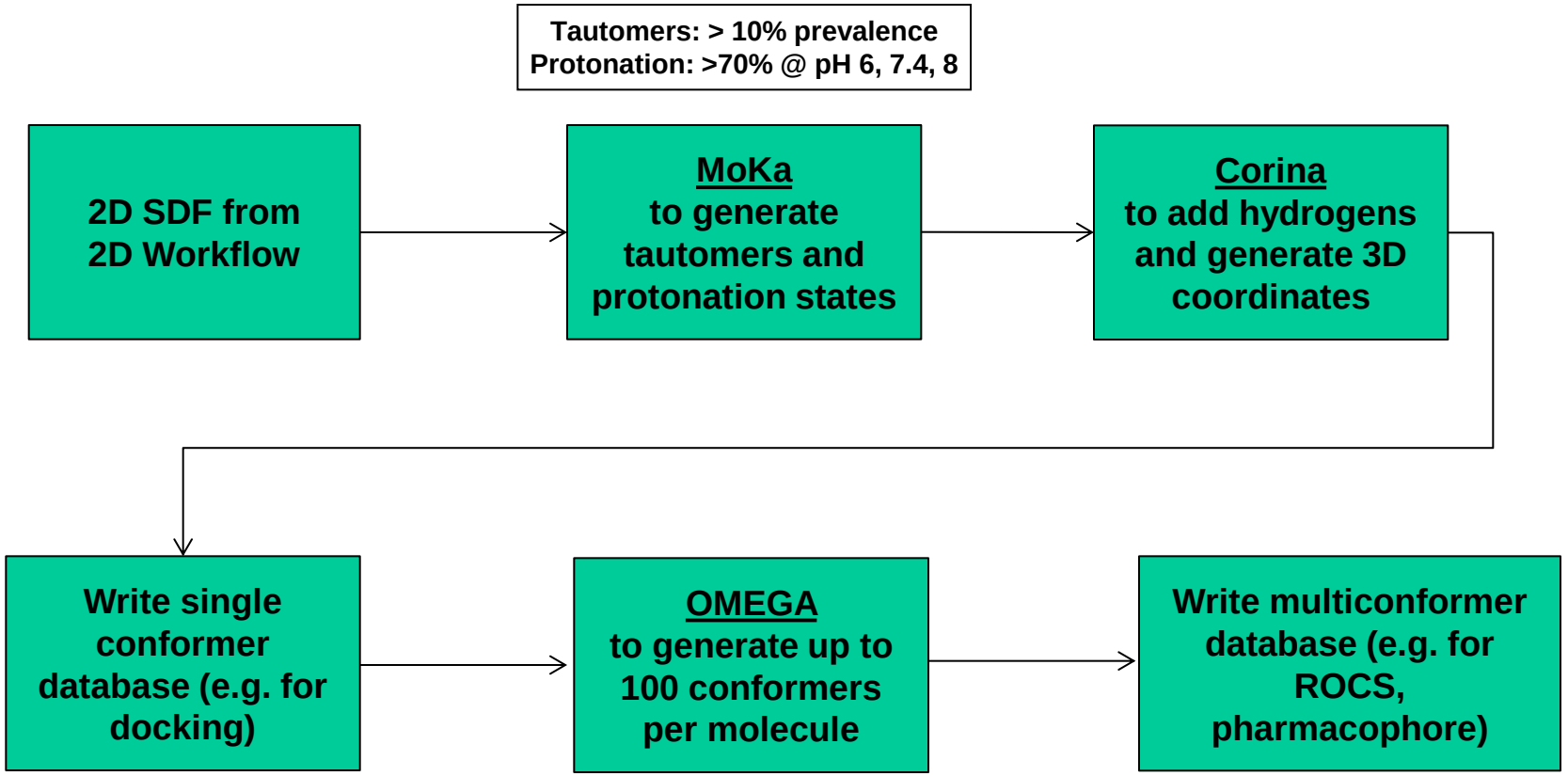
Data Fusion

movit

2D Database Preparation Workflow



3D Database Preparation Workflow



Data Fusion: What is it?

- Combining scores from disparate methods to rank overall results
- Aims to reduce noise by combining results from several methods (presumably “true positives” will rank better by consensus)

JOURNAL OF CHEMICAL INFORMATION AND MODELING

Perspective

pubs.acs.org/jcim

dx.doi.org/10.1021/ci300547g | *J. Chem. Inf. Model.* 2013, 53, 1−10

Combination of Similarity Rankings Using Data Fusion

Peter Willett*

Information School, University of Sheffield, 211 Portobello Street, Sheffield S1 4DP, United Kingdom

JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

dx.doi.org/10.1021/ci300463g | *J. Chem. Inf. Model.* 2013, 53, 1531−1542

Boosting Virtual Screening Enrichments with Data Fusion: Coalescing Hits from Two-Dimensional Fingerprints, Shape, and Docking

G. Madhavi Sastry,[†] V. S. Sandeep Inakollu,[†] and Woody Sherman*[,§]

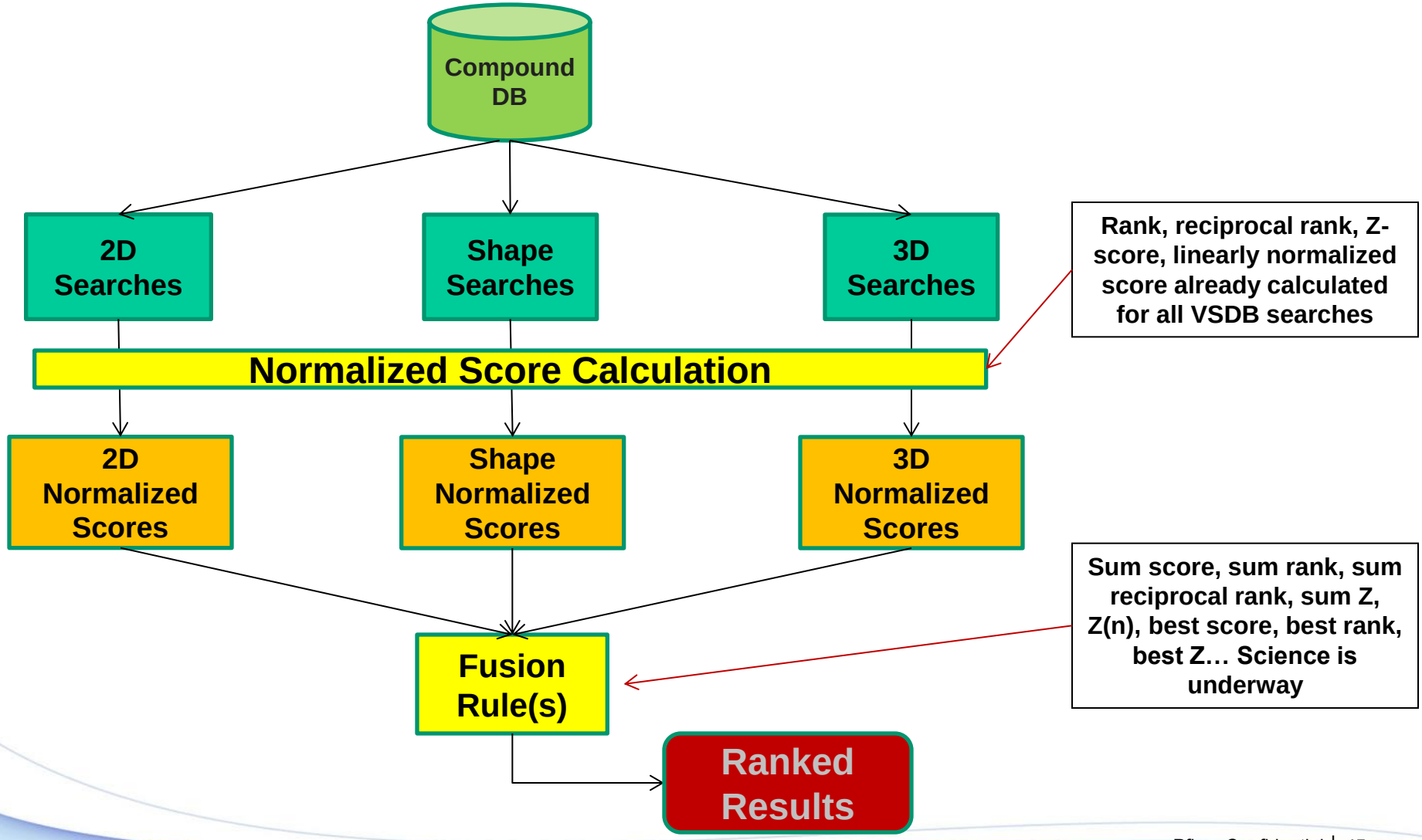
[†]Schrödinger, Sanali Infopark, 8-2-120/113, Banjara Hills, Hyderabad 500034, Andhra Pradesh, India

[§]Schrödinger, 120 West 45th Street, New York, New York 10036, United States

Data Fusion: What Isn't It?

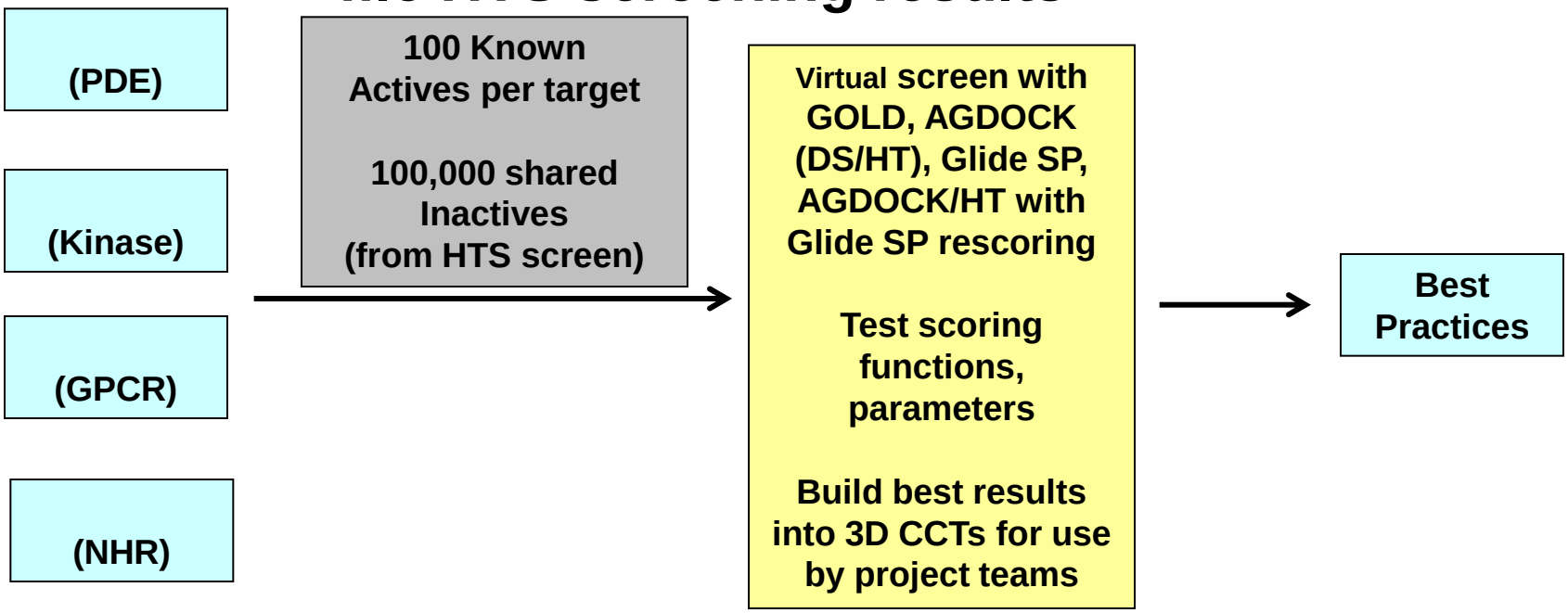
- Data fusion is not a cure-all
  - It will enhance the signal present in existing data, but will not create a signal if it isn't already there
- Data fusion is not a way to combine unrelated data
  - Although you will always get results, this is a garbage-in, garbage-out operation; the better the input data, the more usable the output
- Data fusion is not a way to avoid solid experimental design
  - You have to understand what you are looking for, in order to find it

Data Fusion Workflow



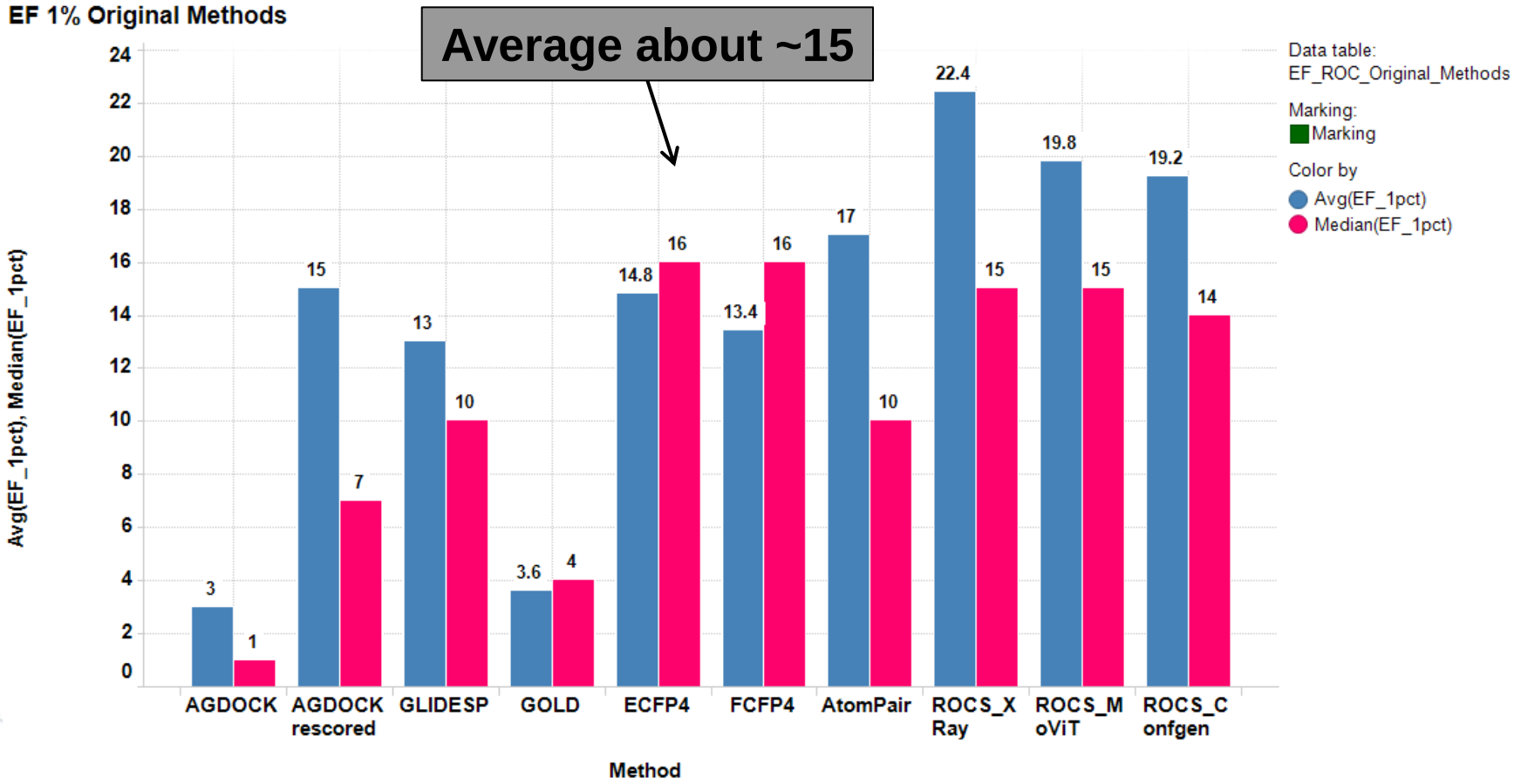
Validation Set

Chose four diverse targets with a crystal structure, validated known actives, and full-file HTS screening results

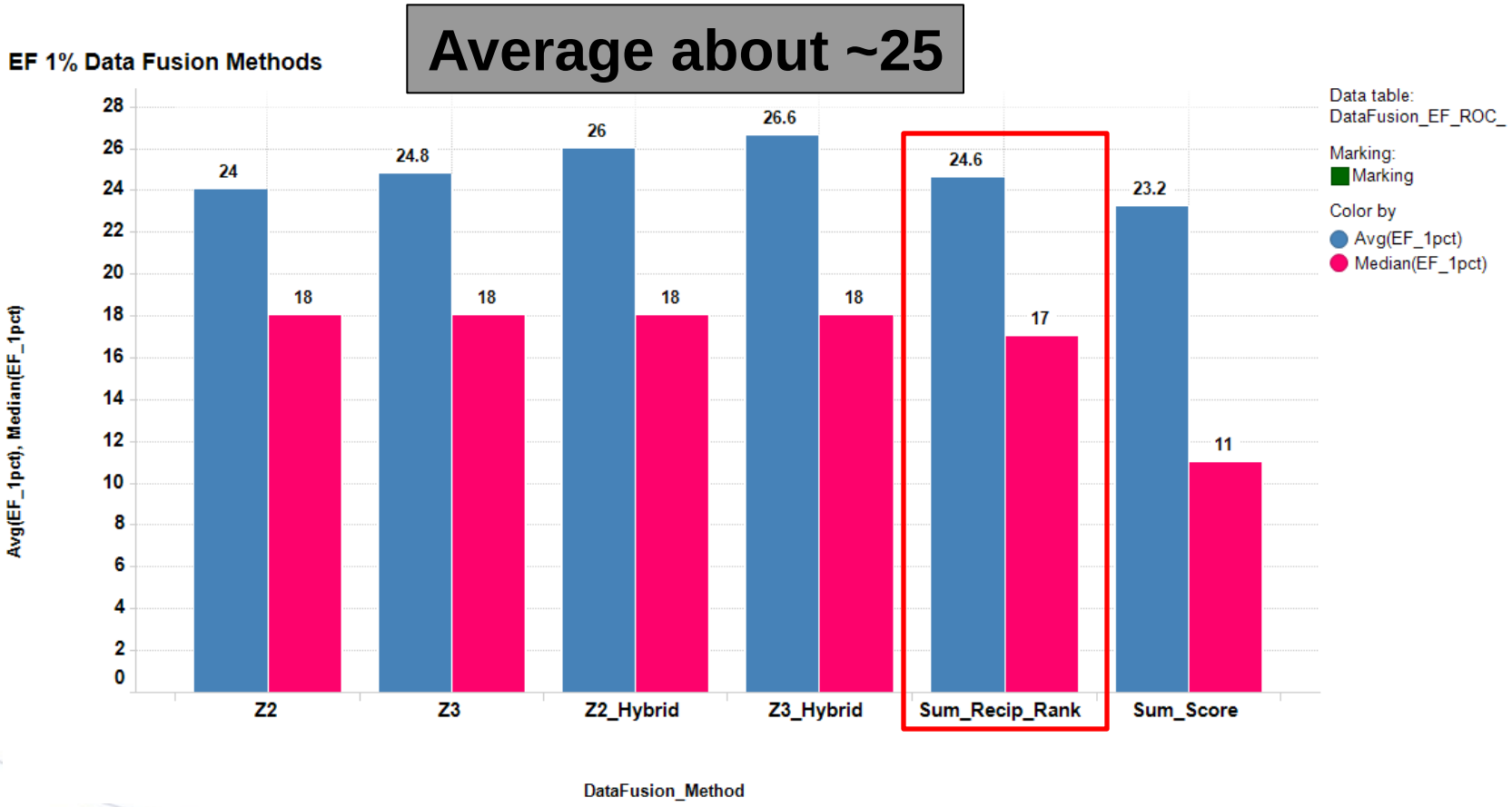


Example: Before Data Fusion

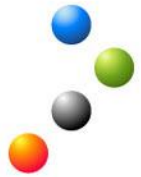
Raw scores from individual virtual screening runs



Results after application of Data Fusion methods



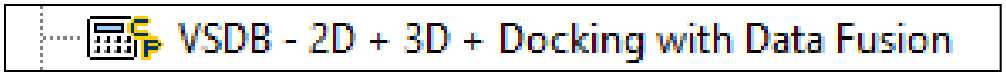
Aside: Which Data Fusion Rule Do I Use?



- Some thoughts:
  - There is little difference among the Z(N) and Sum-Reciprocal-Rank fusion rules for most cases;
  - And… these generally outperform the other methods.

- Why was Sum-Reciprocal-Rank chosen as the default?
  - Reasonable theoretical basis (Willett paper)
  - Uses all of the data, unlike Z(N)

The Protocol of Everything



Property Filtering

Parallel

- 2D methods ← CCT Operands
- FastROCS + ROCS ← 2D/3D probe(s)
- MOE Pharm. ← Pharmacophores

Sequential or Parallel

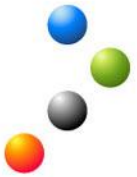
- Protein and defining ligand
- Docking methods

Data Fusion

Experimental Design

- The protocol cannot prevent jobs where 2D, ROCS, pharmacophore, and docking are completely unrelated
- You must keep in mind what your desired endpoints are (e.g. lead discovery versus hit expansion) and plan accordingly
- Data Fusion will rank recurring hits higher – so choose methods that compliment your desired outcome!

Examples



Analog Search

Single probe

↓

AtomPair
PLP
MOE

↓

ROCS with tight cutoff

↓

Sequential Docking

Lead Search

Diverse probes

↓

Feature Trees
Feature Analysis

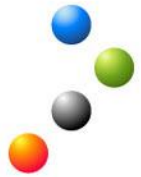
↓

ROCS with loose cutoff

↓

Parallel Docking

Future Directions: Large Scale

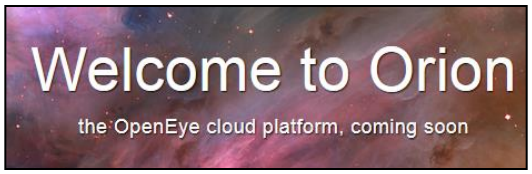


Mattrack
(10^6 cmpds)

Enamine
(10^8 cmpds)

PMC Space
(10^{15-20} cmpds)

Docking: 1 CPU Year



amazon web services

10^5 CPUs / GPUs ($'s!)

Must be clever here!
New algorithms, approaches

BACKUP