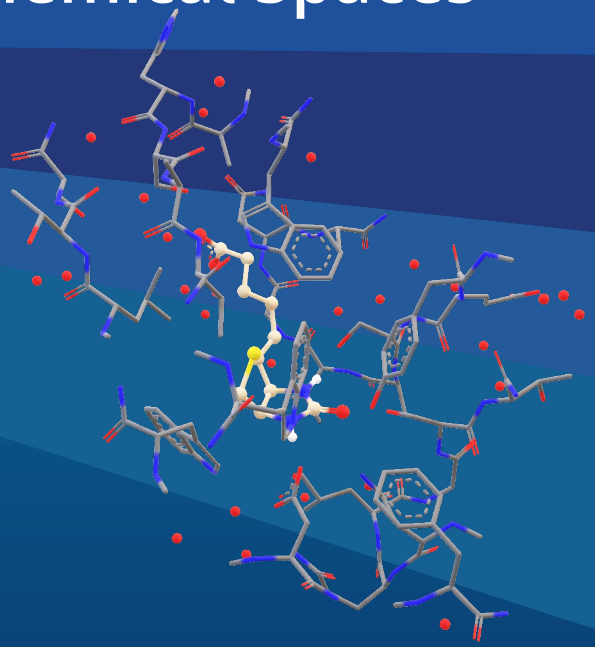
infiniSee

BioSolveIT
expect actives!

Marcus Gastreich

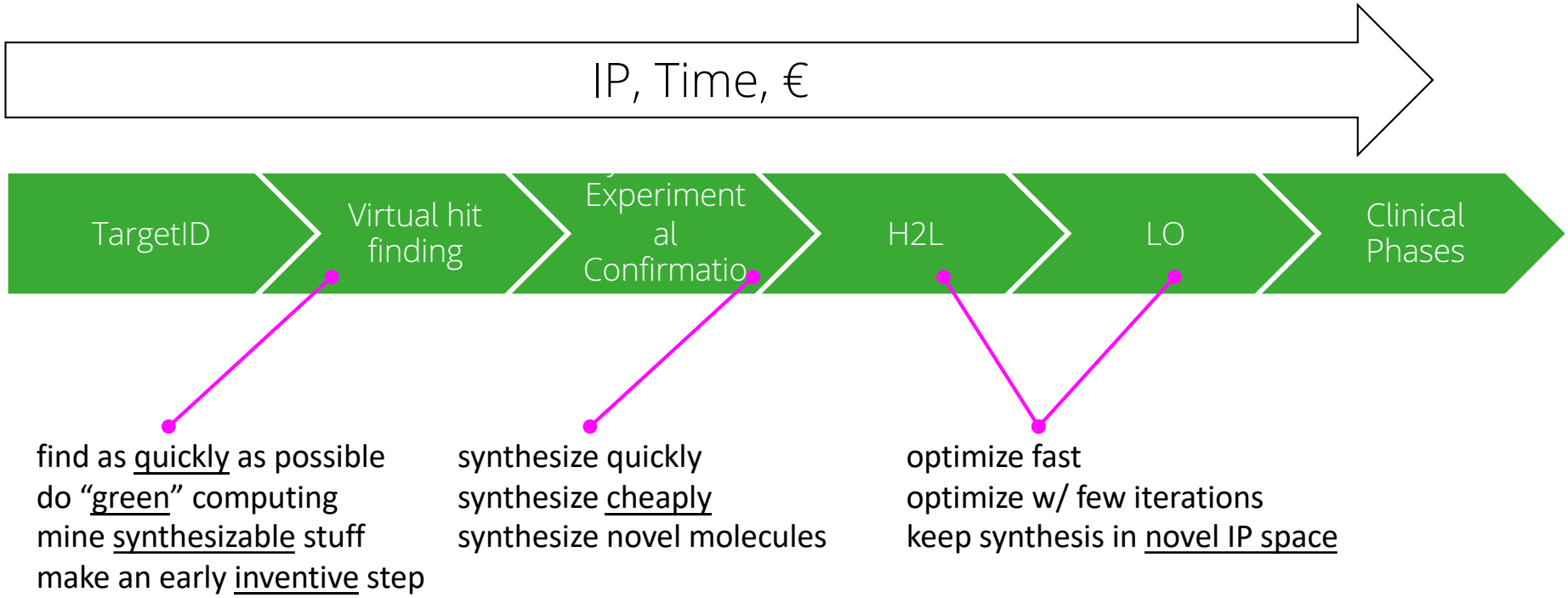
Why Everybody Speaks about Chemical Spaces

A Biased Perspective



Relevance

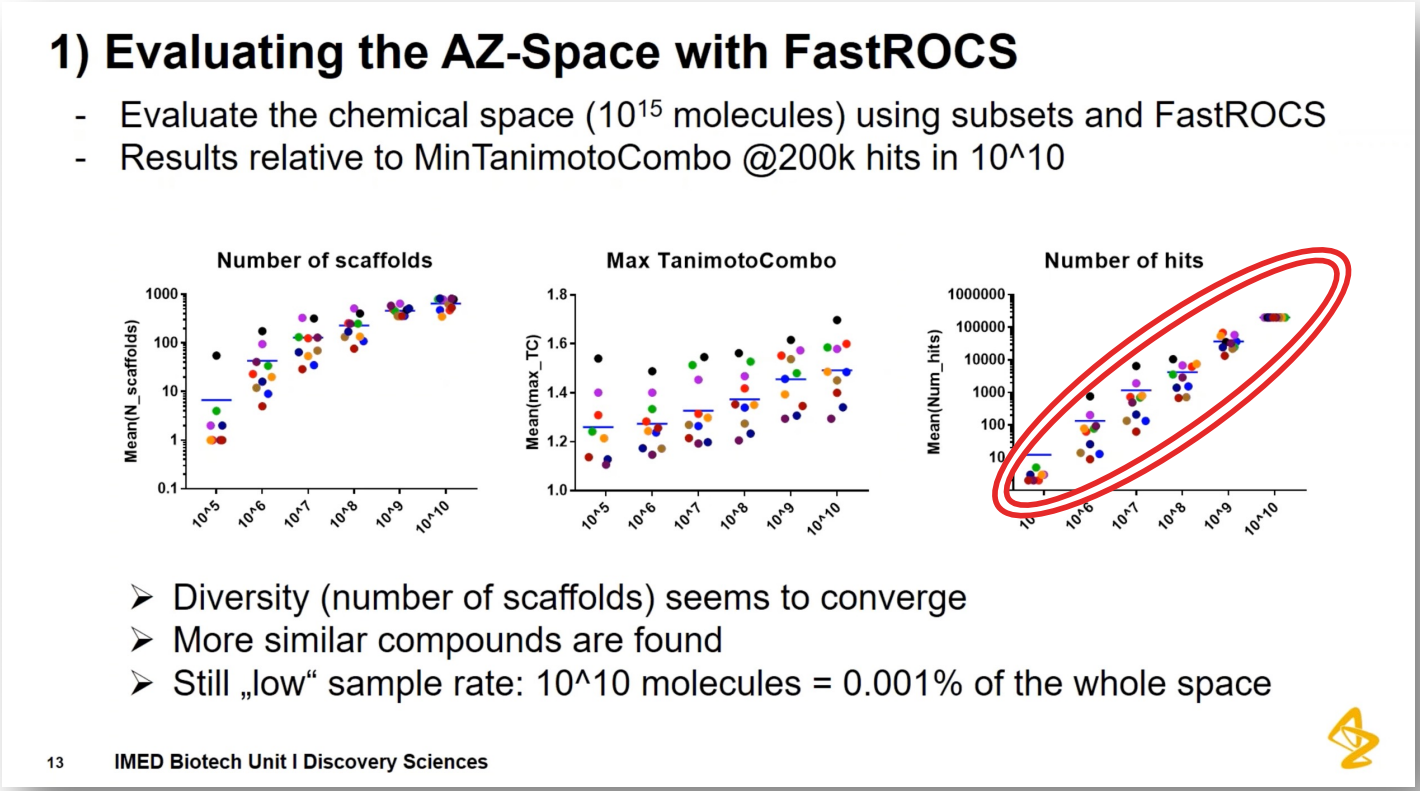
Now that Matthias has explained the methods so beautifully…
Why is all this so relevant — especially in industry?



1 IP: Is Bigger = Better?

Bigger Haystacks -> More Needles

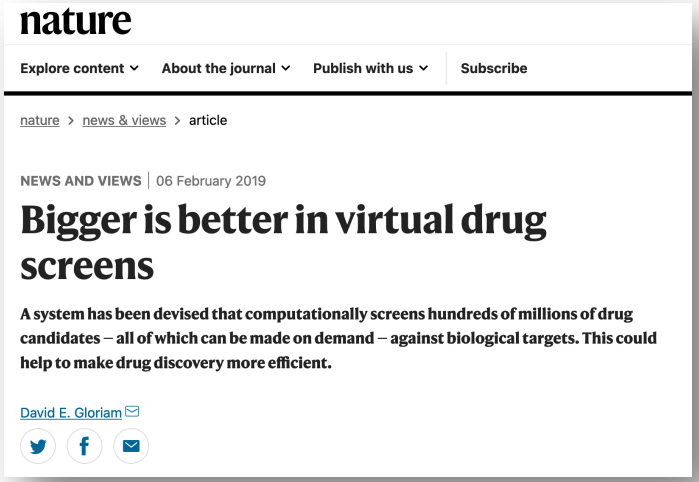
AZ 2018: 10^{10}-sized subsets of **10^{15}**, analyzed with FastROCS



C. Grebner, BioSolveIT Webinar 2018, https://www.youtube.com/watch?v=fMrI11SXwpU

Bigger Haystacks -> More Needles

Lyu et al., Nature 2019, 566, 224: “Bigger is Better”



nature

Explore content · About the journal · Publish with us · Subscribe

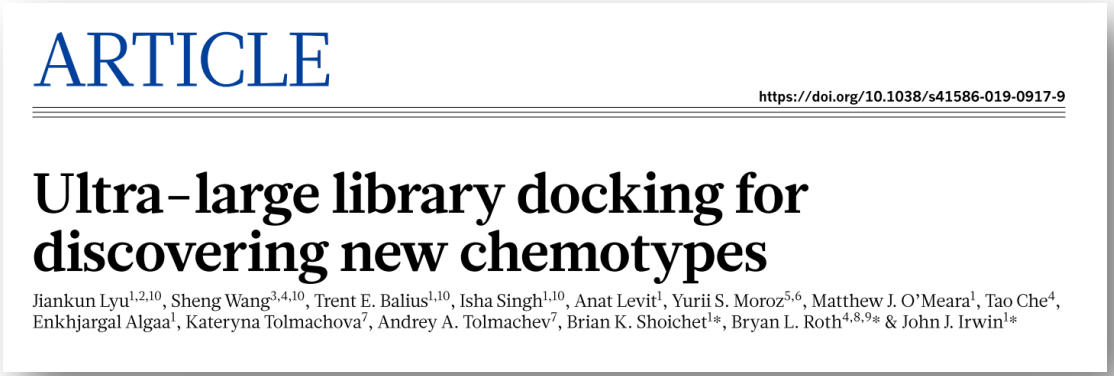
nature › news & views › article

NEWS AND VIEWS | 06 February 2019

Bigger is better in virtual drug screens

A system has been devised that computationally screens hundreds of millions of drug candidates – all of which can be made on demand – against biological targets. This could help to make drug discovery more efficient.

David E. Gloriam

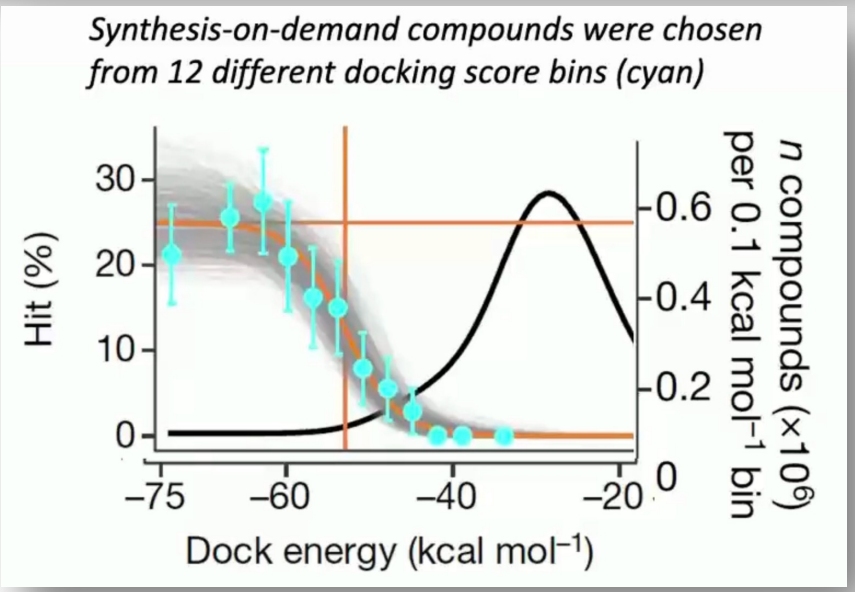


ARTICLE

https://doi.org/10.1038/s41586-019-0917-9

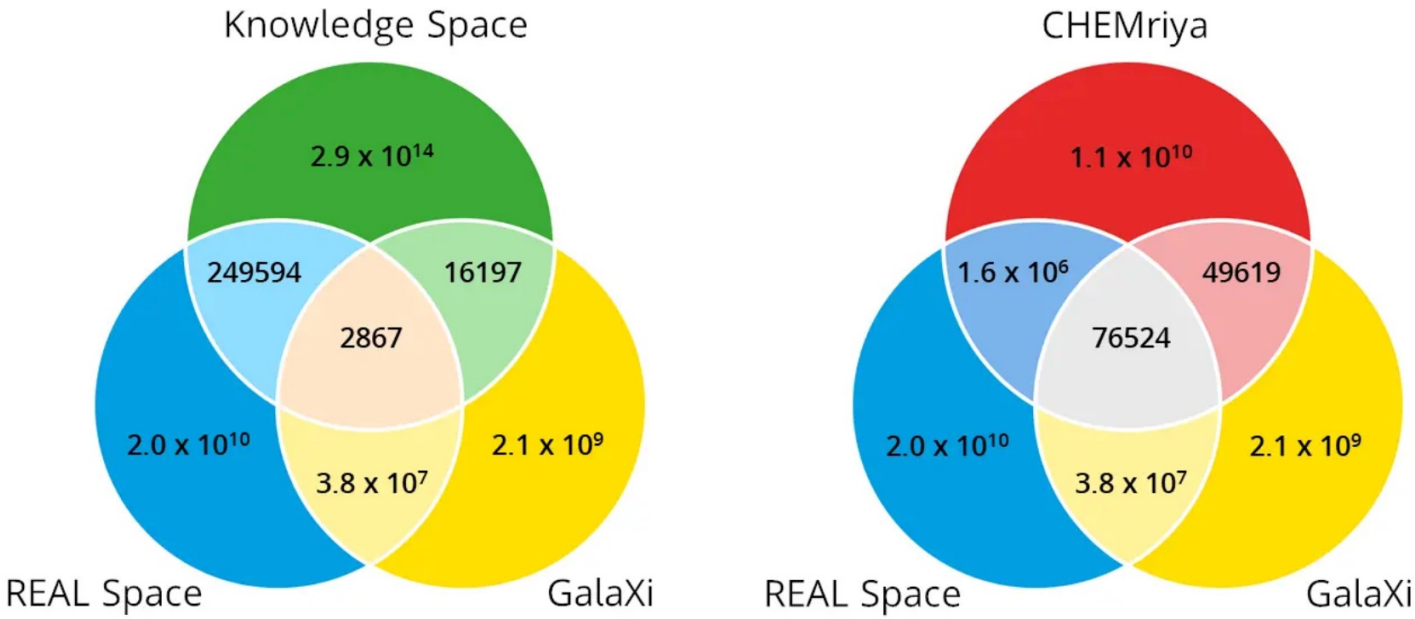
Ultra-large library docking for discovering new chemotypes

Jiankun Lyu[1,2,10], Sheng Wang[3,4,10], Trent E. Balius[1,10], Isha Singh[1,10], Anat Levit[1], Yurii S. Moroz[5,6], Matthew J. O'Meara[1], Tao Che[4], Enkhjargal Algaa[1], Kateryna Tolmachova[7], Andrey A. Tolmachev[7], Brian K. Shoichet[1]*, Bryan L. Roth[4,8,9]* & John J. Irwin[1]*



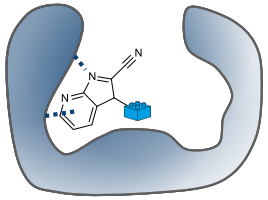
Don‘t Bigger and Bigger Spaces Overlap?

NO !



Bellmann et al., JCIM 2022, 62, 553
so see: Lessel et al, ACS MedChem Letters 2019, 10, 10, 1504-1510

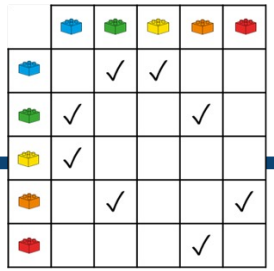
New IP is Found: Extending a Fragment into 3D Space



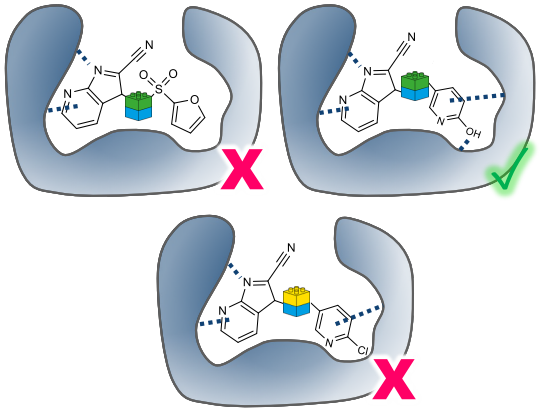
Frag. hit

extend with all building blocks (small % of 10^X)

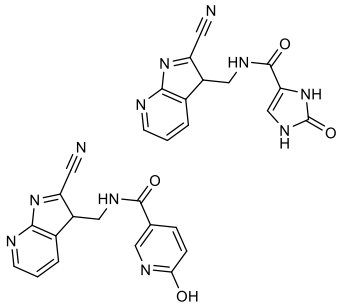
compatibility



vis. & man. filtering



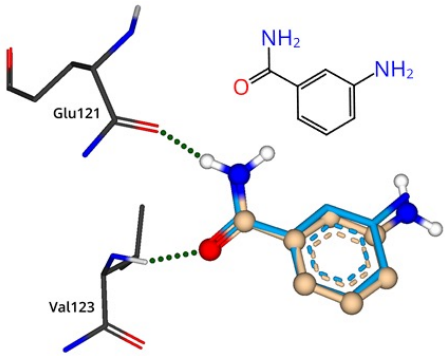
round of filtering



synth'y accessible hits <u>by definition</u>

From 1 Fragment to 69 Purchased New Actives

- Instead from starting at “apo:” Use frag and grow into “Purchasable Space”



x-ray vs. docking: RMSD: 0.3 Å

9 weeks in total

Results:
69 w/ K_i < 500µM
6 x-rays obtained

best binder: 744 nM
(4 rings, 2 stereo cntrs.)

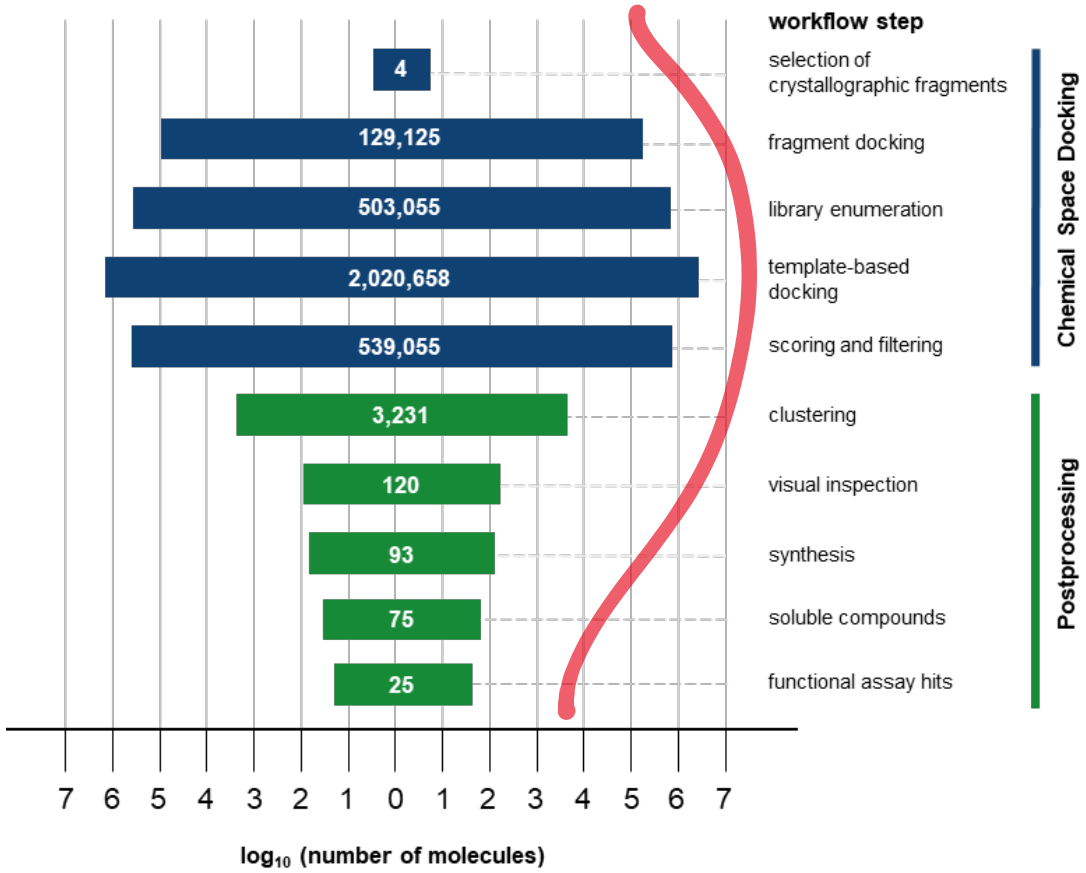
pick stable x-ray → template-dock all Space BBs → score, select → complete the applicable space → template-dock extended fragments → score / cluster / filter / triage

FlexX, fuzzy matching

FlexX, fuzzy matching

Janis Müller (CrystalsFirst) with Enamine, Gerhard Klebe, BioSolveIT, submitted.

Ceci n'est pas un 'funnel'...



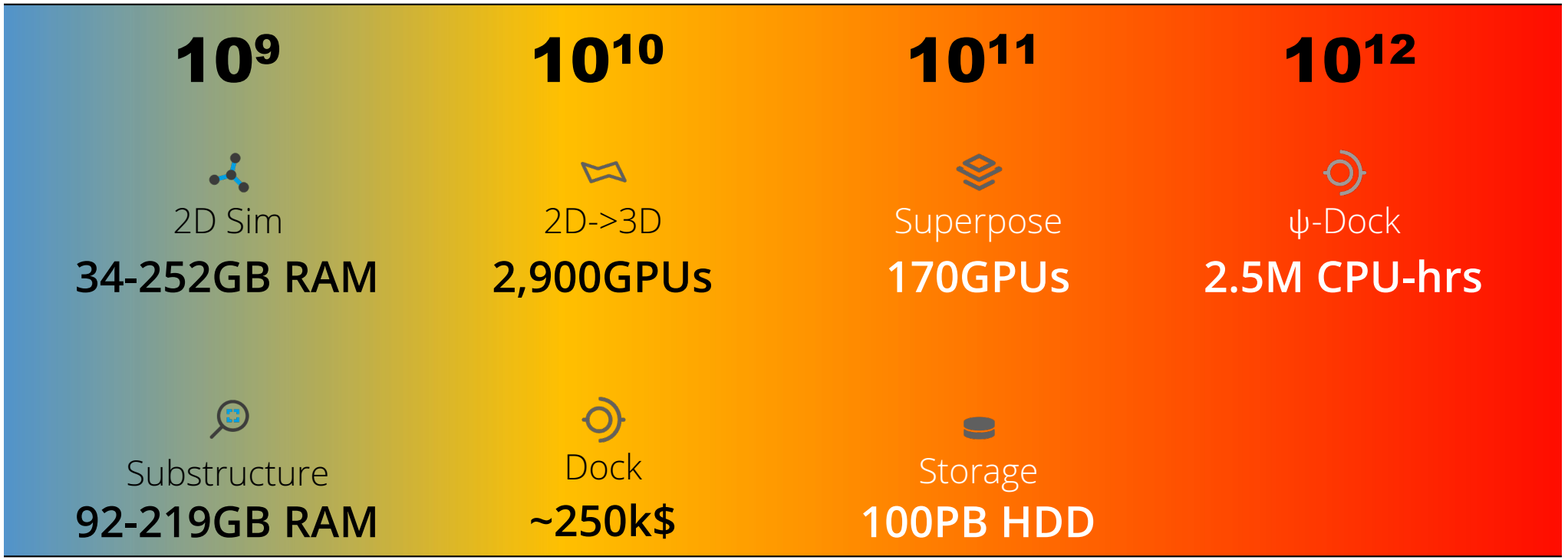
...and, *a priori*, **any** space member may emerge as a hit !

2

€: Combinatorial = Cheaper

A. Virtual
B. In the lab

Enumeration: How to Blow the Budget...



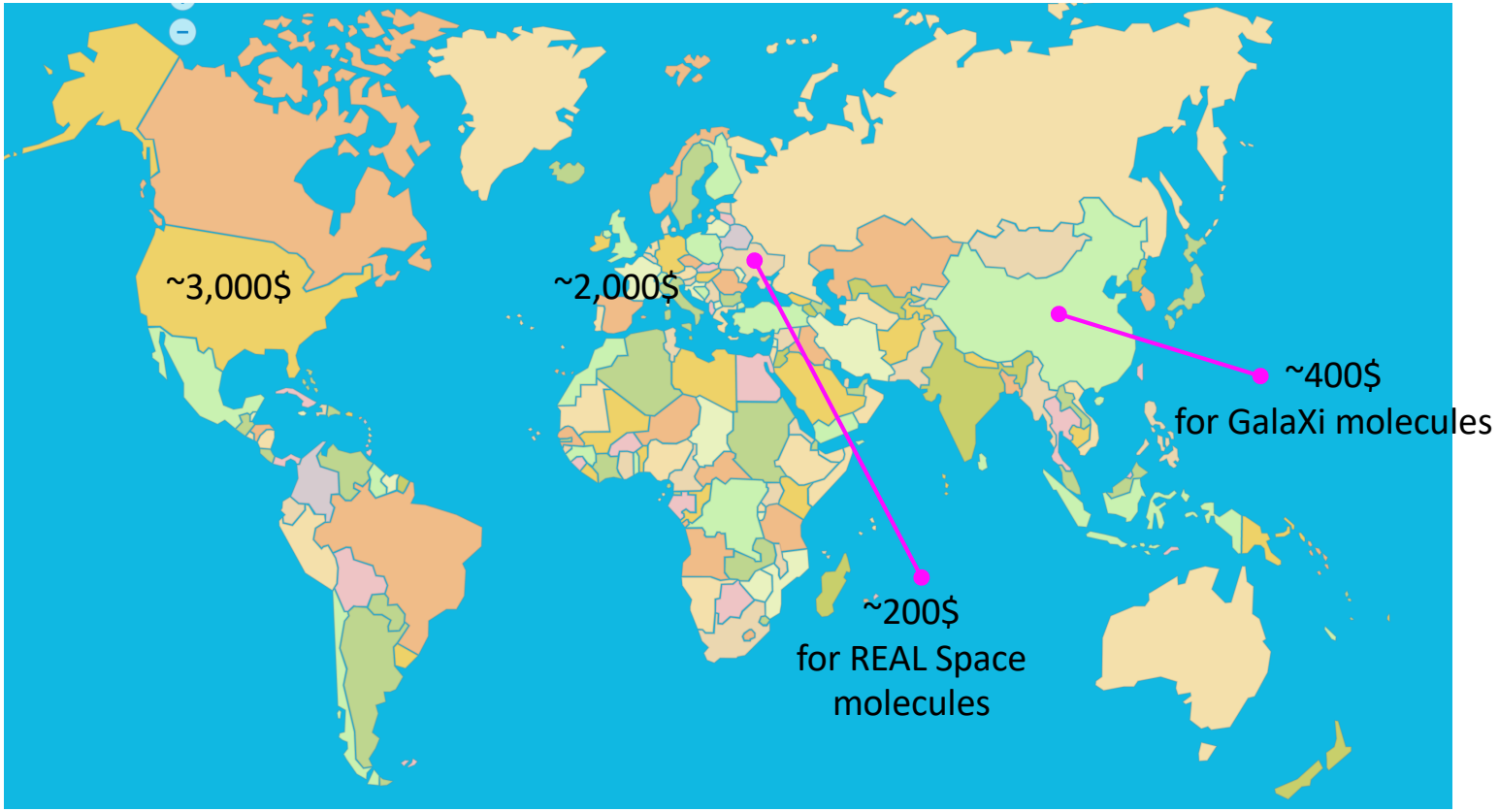
NIH Symposium on Ultra-Large DBs 2020/12, OE also: Grebner C 2018: youtube.com/watch?v=fMrI11SXwpU; also see Review by @WendyAnneWarr, talks (partially): cactus.nci.nih.gov/presentations/NIHBigDB_2020-12

Enumeration: How to Blow the Budget…

Task	Size	CPUs or GPUs	Memory	Cost [$]	time	Ref
Storage	10^9		1(smi)–4(sd) TB			1, NIH
	10^{11}		100PB	1.1M p.a.		1, Google
Gen SMI	10^{10}	400			1.5d	1, OpenEye (OE)
2D Sim	10^9				9h	1, molsoft
	10^9		34–252GB			1, NextMove
Conformers	10^{10}	2.900	20TB	20K		1, OE
3D Superposition	10^{11}	170		300	58m	1, 3 OE/AZ
Docking	10^{10}		~100TB (est.)	~250.000		1, OE
	10^9	1.500			1.5d	4
Surrogate Docking	10^{12}	"1"		6M amort.	2.500.000h	1, Argonne
Substructure	10^9		92–219GB			1, NextMove

1) NIH Symposium on Ultra-Large DBs 2020/12, OE also: Grebner C 2018: youtube.com/watch?v=fMrI11SXwpU
see pending Review by @WendyAnneWarr, talks (partially): cactus.nci.nih.gov/presentations/NIHBigDB_2020-12
2) Clark, D, JCIM 2020, DOI: 10.121/acs.jcim.0c00101
3) Grebner et al., JCIM 2019, DOI 10.1021/acs.jcim.b00779 4) Lyu et al., Nature, 2019, DOI 10.1038/s41586-019-0917-9

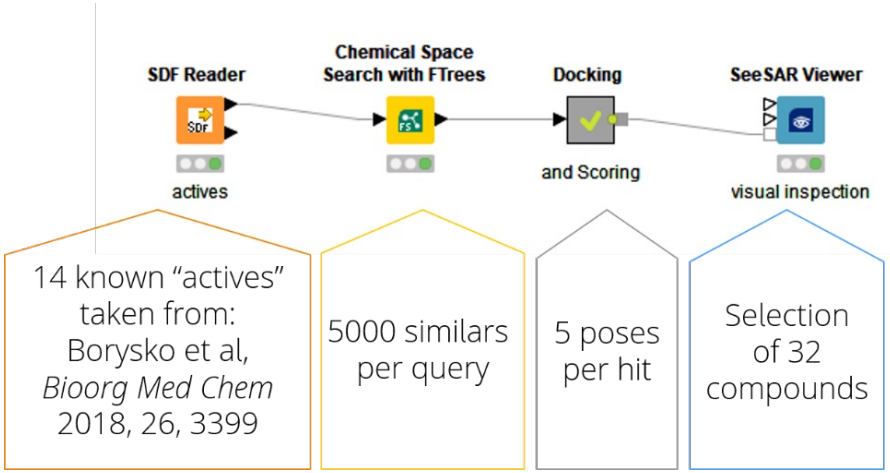
In the Lab: Synthetic Costs per Molecule

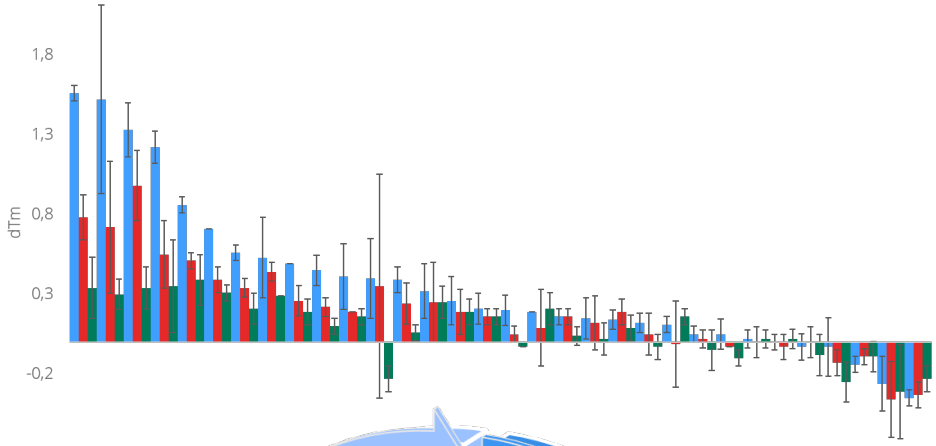


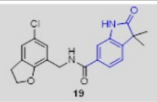
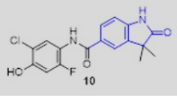
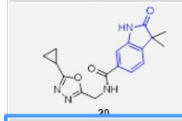
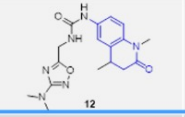
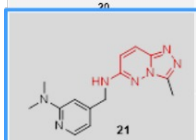
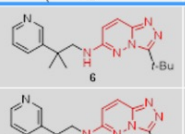
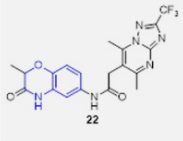
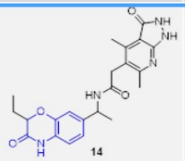
Coarse values, by hearsay

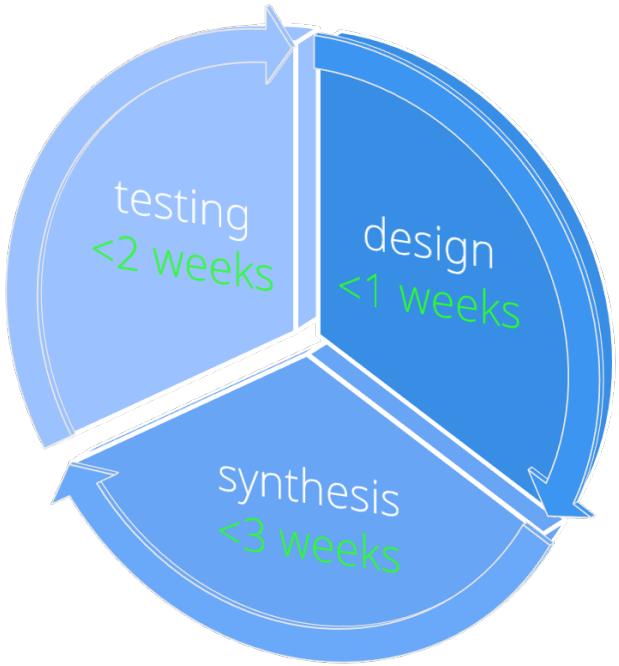
3 Time: Combinatorial = Faster

6 Weeks: SAR-bySpace Around BRD4 Inhibitors





Query	Hit	Tanimoto sim	FTrees sim	ΔT_m (@40µM)	IC_{50} (µM)
		0.456	0.956	0.7	10
		0.277	0.920	1.5	26
		0.323	0.933	1.3	44
		0.333	0.953	1.5	68
		0.356	0.932	0.5	141



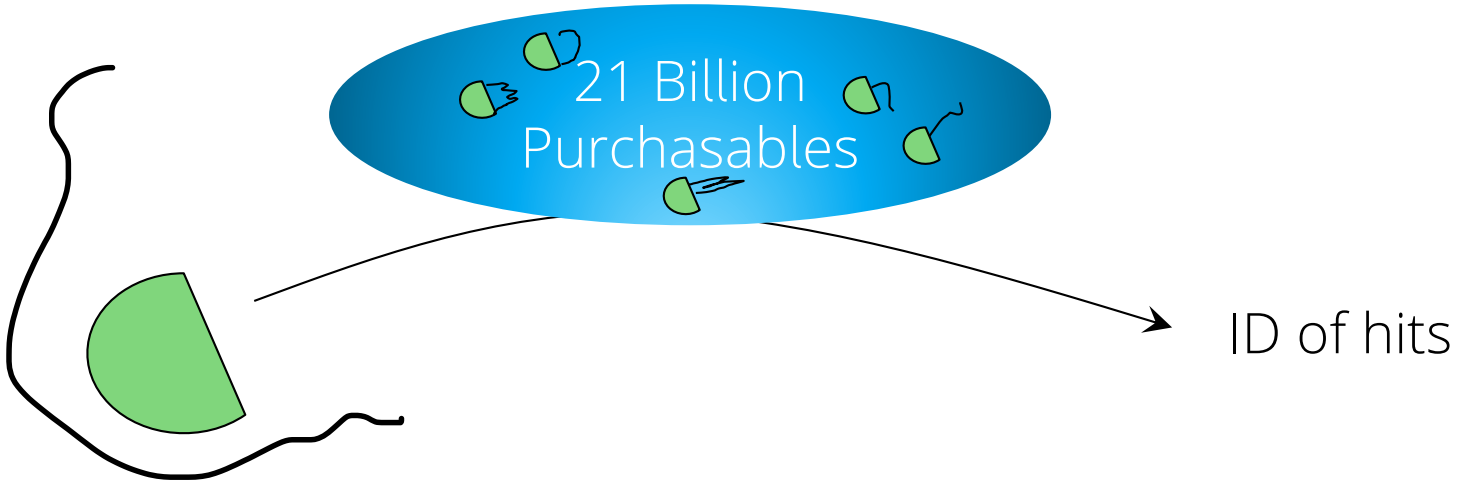
*Klingler et al., Molecules 2019, 24(17), 3096

Xavi Barril @DrugSpace Symposium 2022

Howto: Searching the purchasable space in 3D ??

Xavi's classical / traditional way

Idea: Search substructures in vast space, place, test:



X. Barril: Fragment Extension by Matching Substructures

SpaceMACS pulls out molecules from giga…zetta spaces

SpaceMACS: Connecting Top and Bottom

Very fast!

- 20M query: 1 - 400 CPU hours (parallelizable)
- Big disparity depending on query
- Big outputs = Large Memory

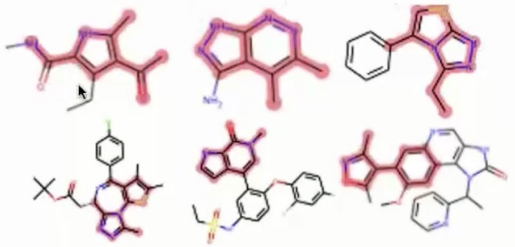
```
[Result]
NofResults = 20000000
MinimumMappingSize = 10
MinimumSimilarity = -1.0
MinimumResultSize = 25
MaximumResultSize = 35
ShowMappingInsertion = False

[Matching]
Mode = Size
RingChainAtomCompatibility = False
SkipSingleFragmentMCSResults = False

[Limits]
IntermediateResultStorageLimit = 10000000
MaxNofOpenLinkExtensions = 1000000
```

What do you find?

ID	Exact MCS
4LZS	1.055.698
6ZED	14.531
6ZF9	-
ABBV-075	-
IBET151	4.660.449
JQ1	-



Enamine
REAL Space™ 2.1x10[10]

infiniSee

« 1000 times faster than brute force docking »

21Bn → Pull Mios SpaceMACS → Filter → 2D > 3D → Tethered Docking → Cluster → MMGBSA → DUck

X. Barril: Speedup in Synthesis by Purchasable Space Orders

REAL Space: These guys deliver!

	Requested	Obtained
No. Compounds	Round 1: 58 Round 2: 37 Total = 95 (6 discarded as too expensive)	50 (+3) 32 85 (89%)
Purity	>90%	97.7% (avg) all > 95%
Amount	10mg	9.7mg (avg) 86% >9mg
Delivery time		5 weeks (August) 4 weeks (November)

Yurii Moroz

Inna Khutorianska

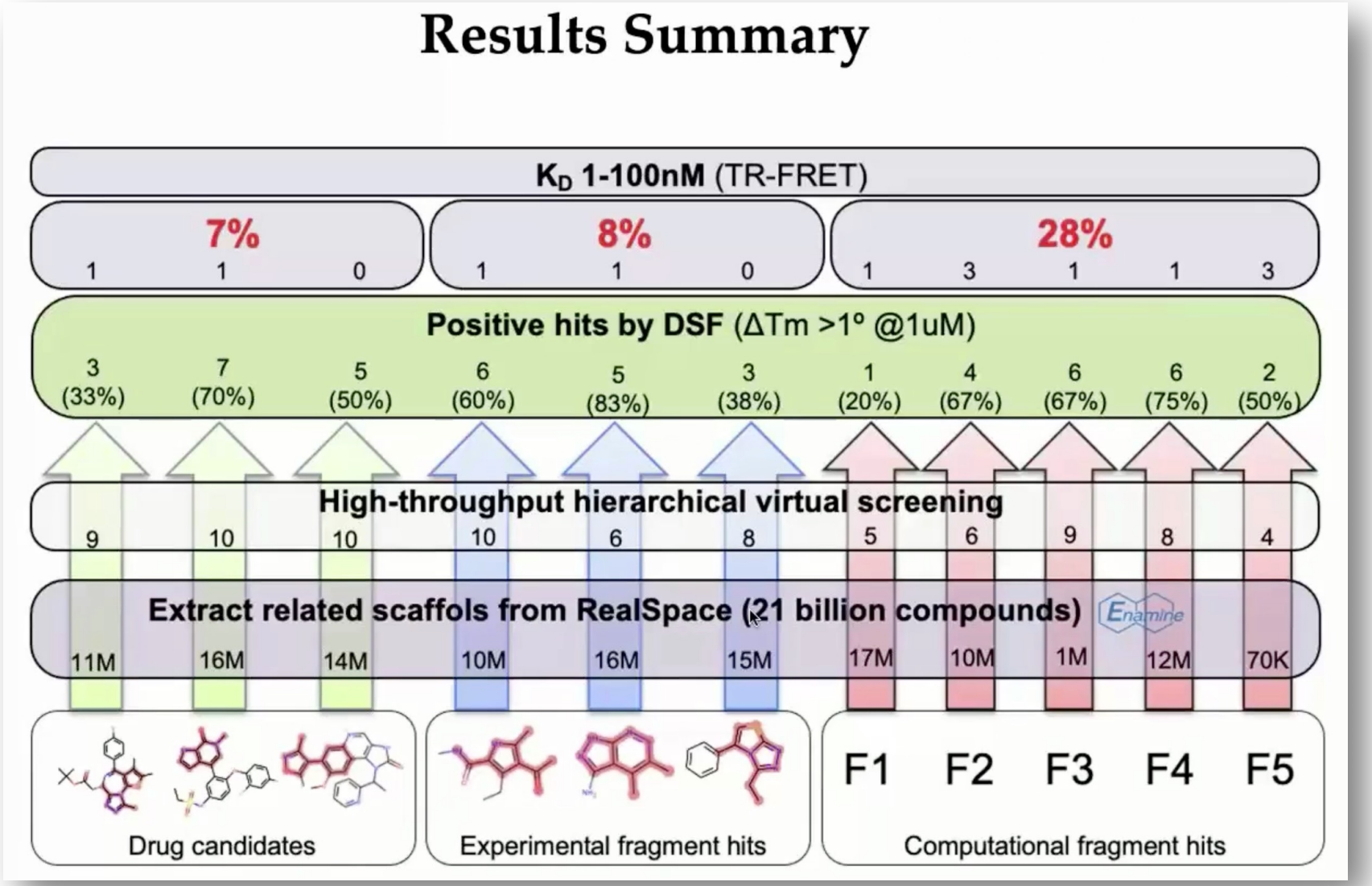
https://enamine.net/news-events/1336-call-for-donations

We in Ukraine extremely need now strong support of the world community. Please consider donation to Enamine's charity fund

Enamine

REAL Space™ 2.1x10^{10}

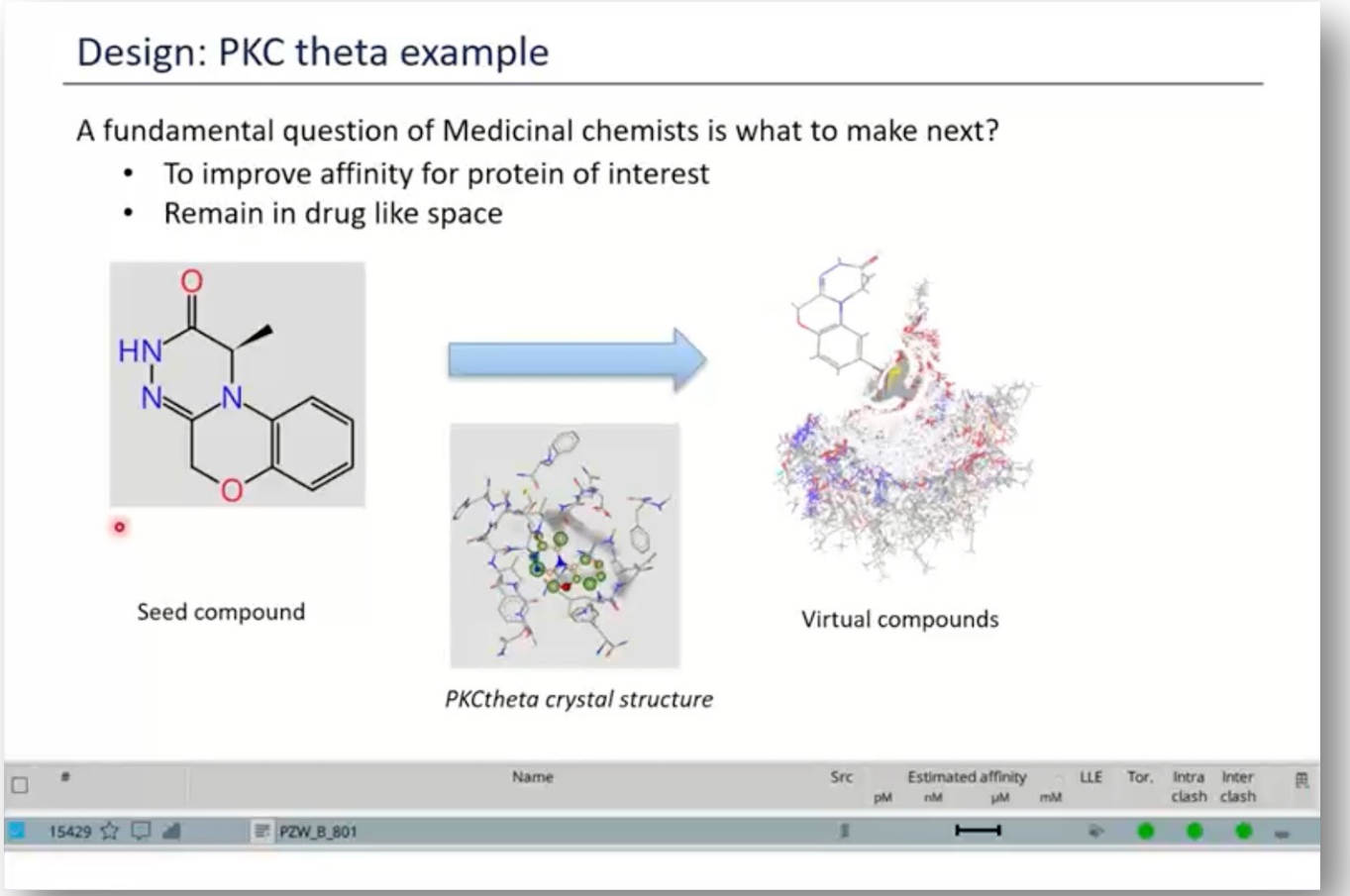
X. Barril @DrugSpace22



AbbVie's Jeremy Edmunds @DrugSpace Symposium 2022

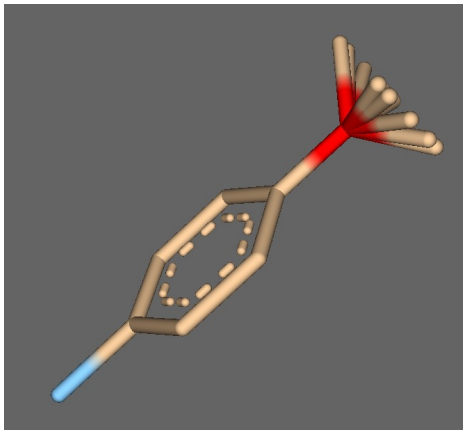
Use new FastGrow app to dramatically speedup candidate generation!

Idea: Grow systematically from all sensible exit vectors,
Use purchasable building blocks:

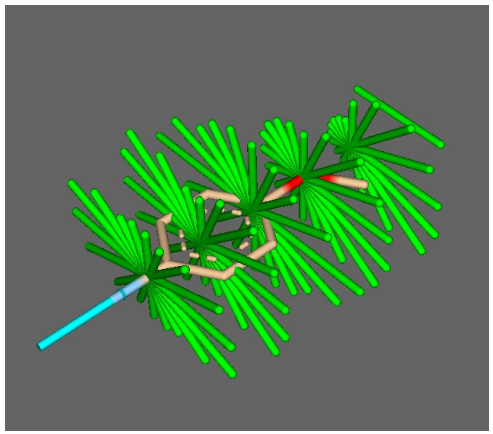


Superfast Growing with FastGrow

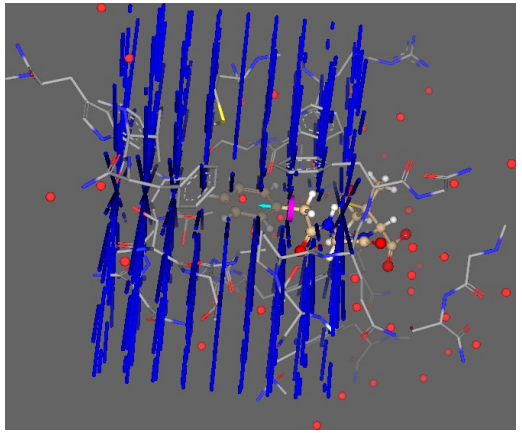
Compatible Compound Selection → Shape Matching → Sphere Constraint Filter I → Reactive SMARTS Filter → Restrained Opt → Clash Filter → Sphere Constraint Filter II → *n* Compounds



Conformer sampling



Conformer descriptor
(1 conformer)



Pocket descriptor

- Grows from **12k fragments in ~10 seconds using bit shifting** $\hat{}$ **in lieu of** R
- Collaboration with Servier Paris and ZBH (Patrick Penner @Rarey Lab)
- Manuscript submitted

AbbVie @DrugSpace Symposium 2022

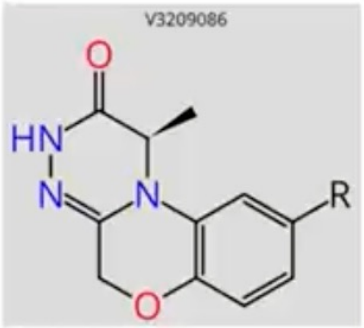
FastGrow 0.9.3

~ 3.6 million fragments : 31 GB

Combined200.fastgrow

2h 20 min

Limit 10,000 compounds



~ 3.6 million fragments : 31 GB

Combined200.fastgrow

3h 40 min

Limit 100,000 compounds

Notice that the amino azetidine appears in the 100,000-selection set, not the 10,000 - selection

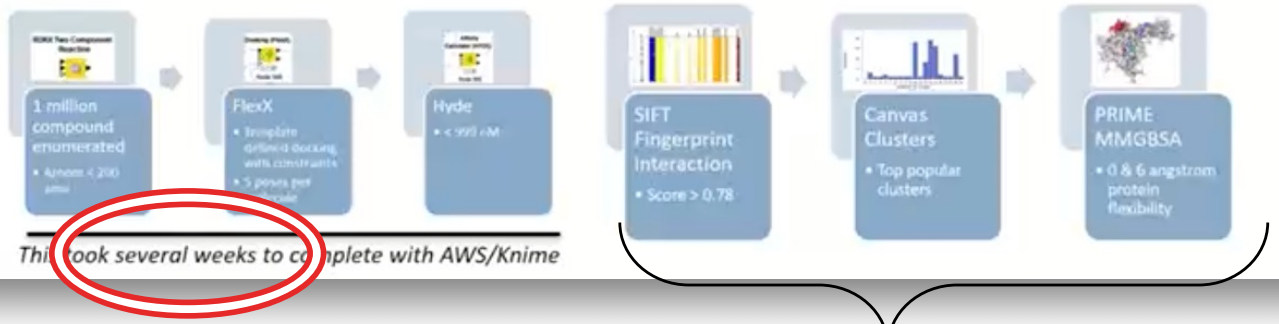
Subsequent Hyde scoring takes approx. 20 mins per 10000 compounds (Knime)

AbbVie @DrugSpace Symposium 2022

before

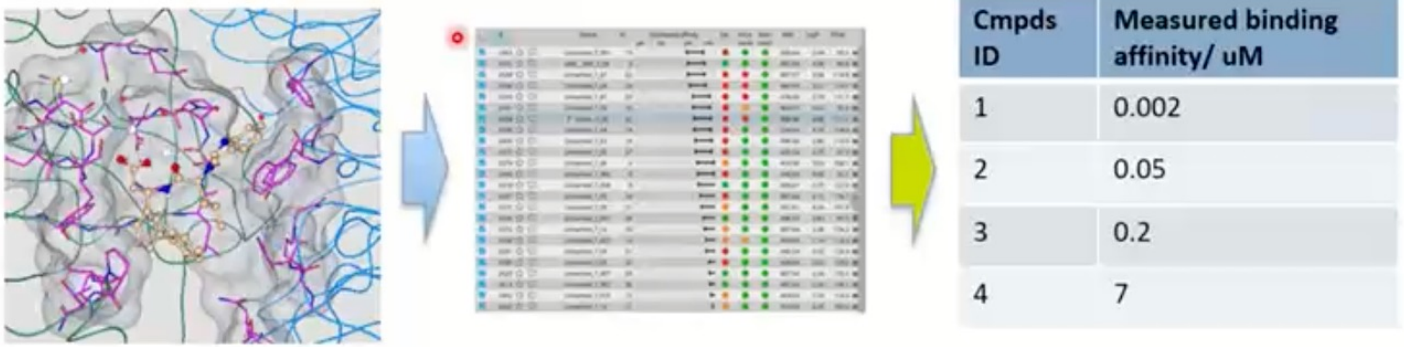
Different project: From about 1 million virtual compound to 4 compounds synthesized

De novo design of compounds using virtual compounds created from A-room fragments, FlexX docking, H Bond/dehydration energy filtering, Interaction map scoring, Clustering for similarity, and finally calculation of ligand binding/ligand strain energies (MMGBSA), and synthesis assessment



now

A few hours



Cmpds ID	Measured binding affinity/ uM
1	0.002
2	0.05
3	0.2
4	7

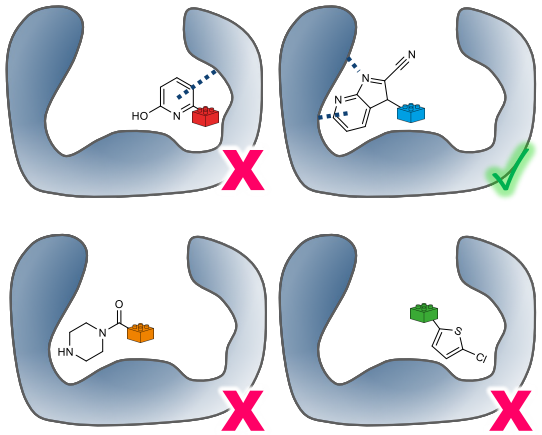
4 very actives

- Selected 4 compounds based on concordance of MMGBSA & Hyde for synthesis and prepared at a CRO

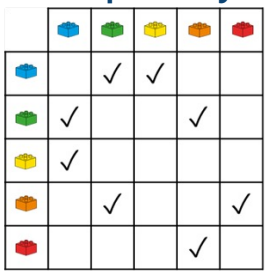
Docking Untruncated Spaces

SPACE
10^X mols
synth'y accessible

docking all building blocks (small % of 10^X)

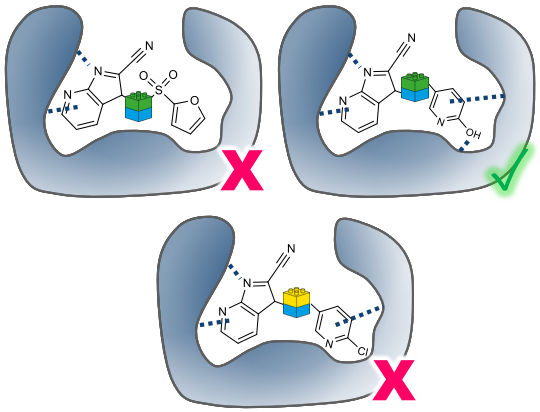


compatibility

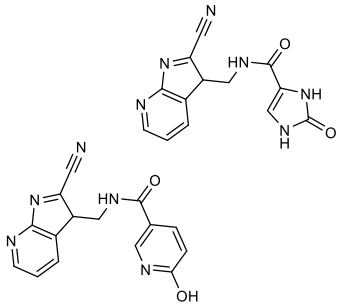


anchor & grow

vis. & man. filtering

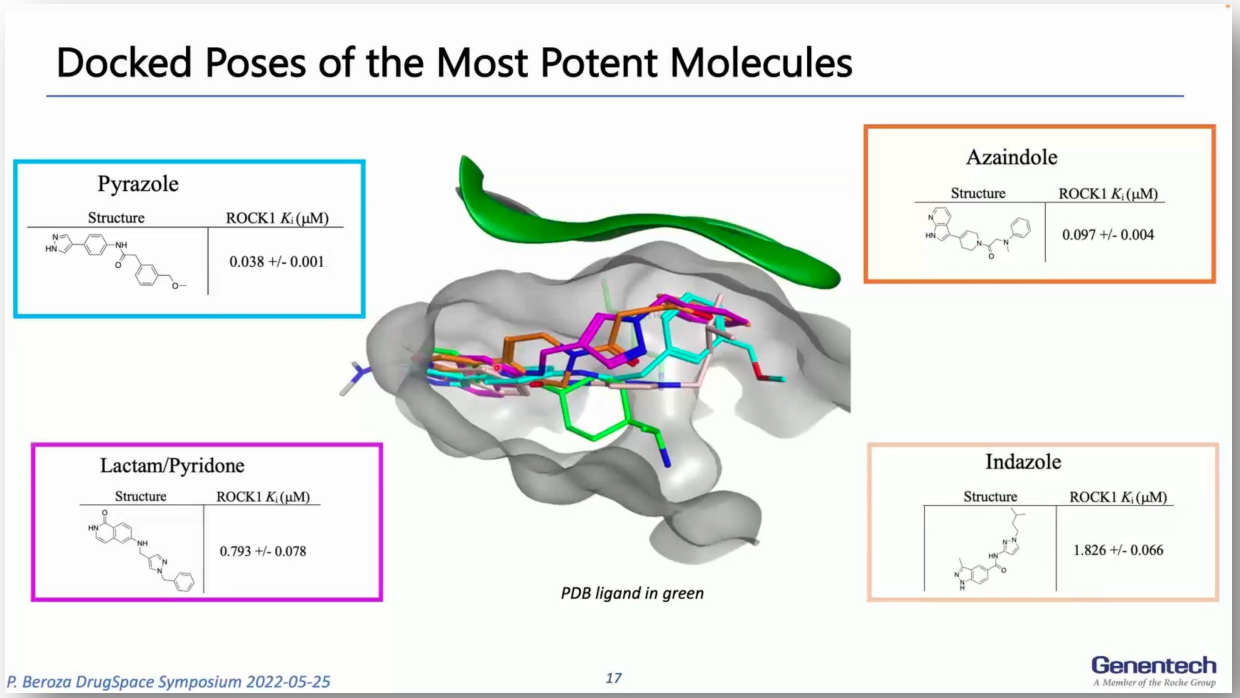


3rd round of filtering



synth'y accessible hits <u>by definition</u>

Genentech & BioSolveIT: Chemical Space Docking

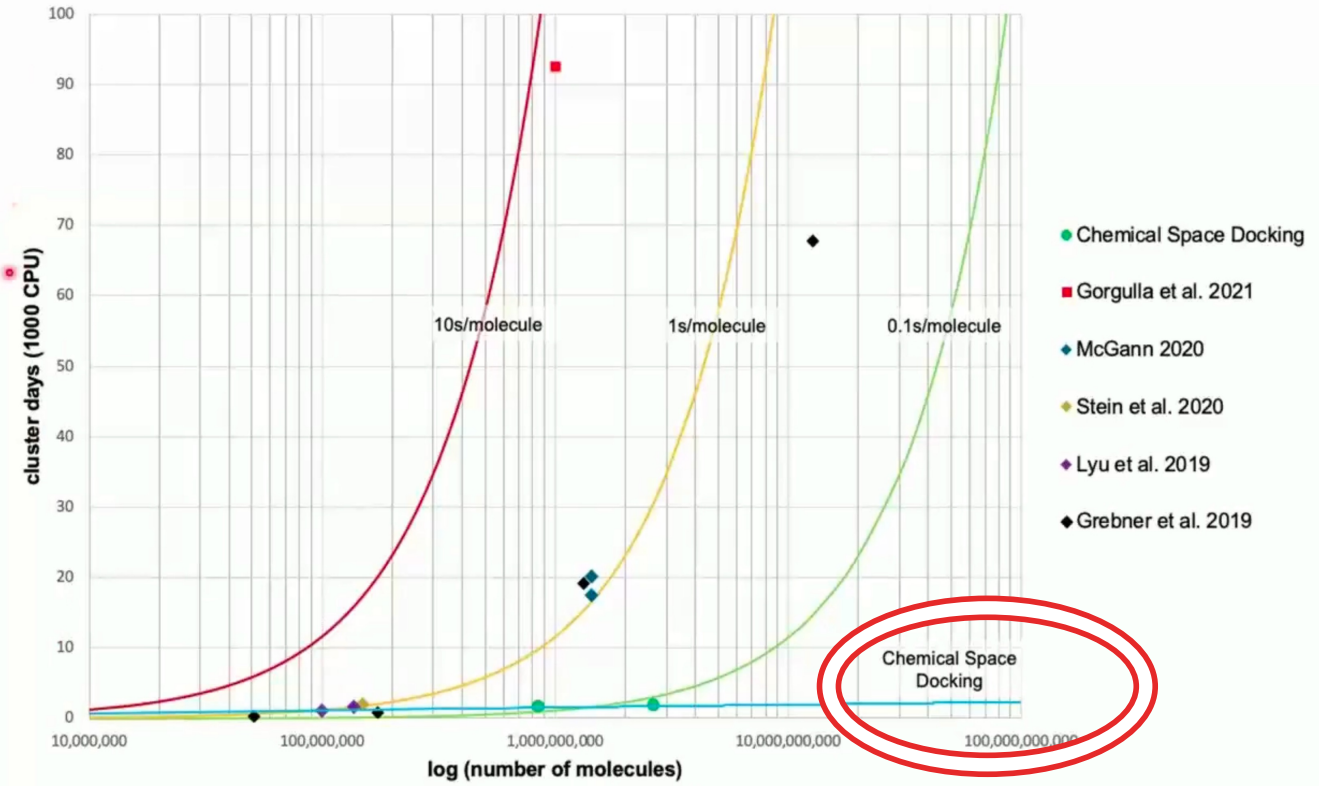


Activity (K_i)	#
10-100 nM	7
0.1 - 1.0 uM	6
1.0 - 10.0 uM	14
>20 uM	42

~40 % hit rate
Most potent: **38 nM**

Chemical Space Docking is Green Computing (in review)

Efficiency of Chemical Space Docking

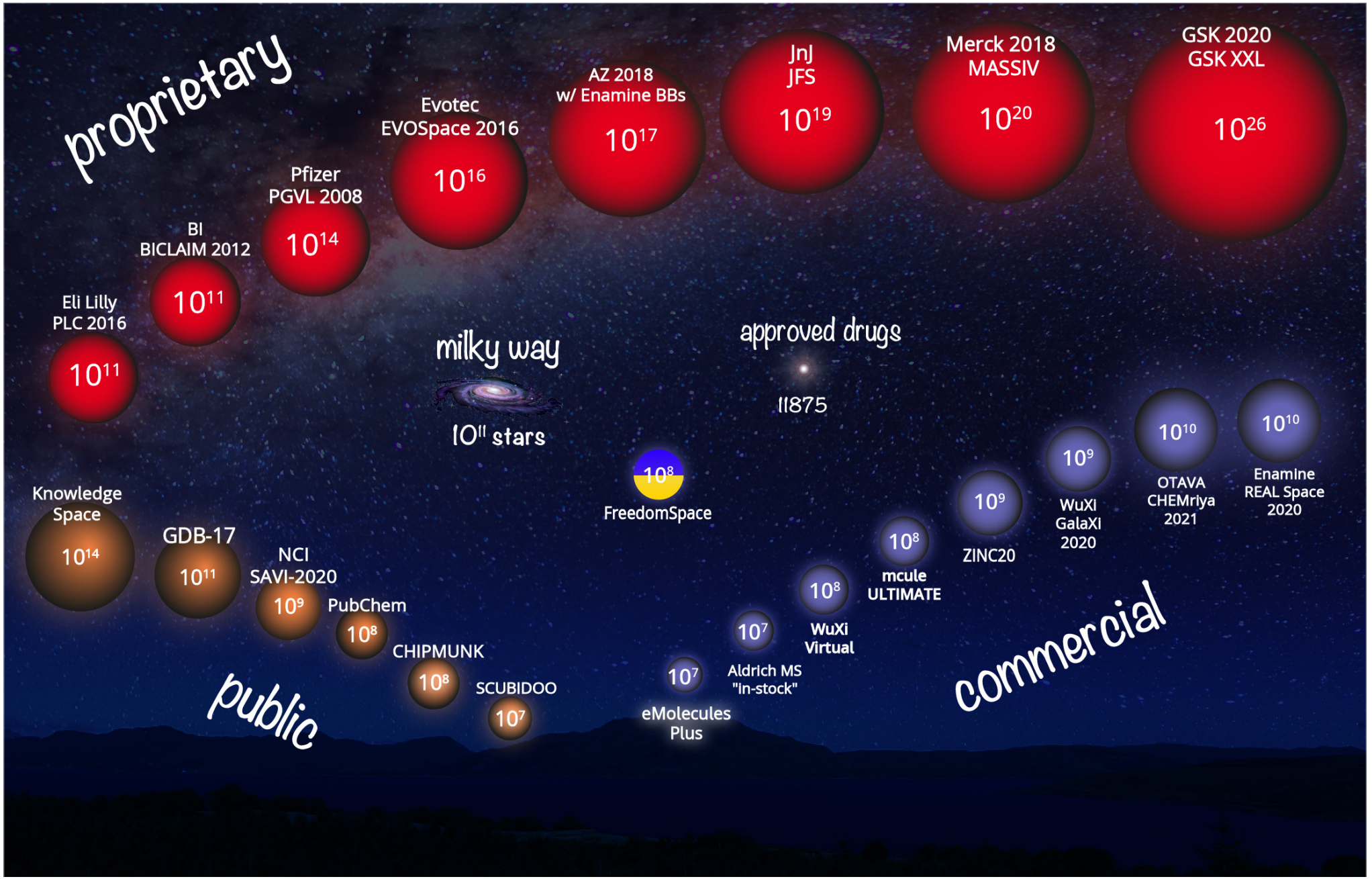


P. Beroza DrugSpace Symposium 2022-05-25

24

Genentech
A Member of the Roche Group

Thank You!


proprietary
Evotec
EVOSpace 2016
10^{16}
AZ 2018
w/ Enamine BBs
10^{17}
JnJ
JFS
10^{19}
Merck 2018
MASSIV
10^{20}
GSK 2020
GSK XXL
10^{26}
Pfizer
PGVL 2008
10^{14}
BI
BICLAIM 2012
10^{11}
Eli Lilly
PLC 2016
10^{11}
milky way
10^{11} stars
approved drugs
11875
10^{8}
FreedomSpace
10^{10}
OTAVA
CHEMriya
2021
10^{10}
Enamine
REAL Space
2020
10^{9}
WuXi
GalaXi
2020
10^{9}
ZINC20
10^{8}
mcule
ULTIMATE
10^{8}
WuXi
Virtual
10^{7}
Aldrich MS
"In-stock"
10^{7}
eMolecules
Plus
commercial
Knowledge
Space
10^{14}
GDB-17
10^{11}
NCI
SAVI-2020
10^{9}
PubChem
10^{8}
CHIPMUNK
10^{8}
SCUBIDOO
10^{7}
public

Join the LinkedIn **Chemical Space Club**!

- Keeping in touch with 400+ global key players
- Discussions & presentations
- Early announcements, paper hints etc.
- linkedin.com/groups/9004052/