pharmit

virtual screening in the cloud with pharmit

David Ryan Koes

@david_koes

Computational and Systems Biology
University of Pittsburgh

The Cloud

The Cloud

Running your stuff on other people's computers.

The Cloud


Trace
Logging
Error Reporting
Bigtable
SQL
Datastore
Storage
Spanner
Monitoring
Debug
Container Engine
Cloud Functions
Networking
Running your stuff on other people's computers.
BigQuery
Pub/Sub
App Engine
Compute Engine
ML Engine
Genomics
Container Registry
Development
Dataproc
Dataflow

Why Cloud?

To get access to resources.

- compute power
- software
- storage
- data

Why Cloud?

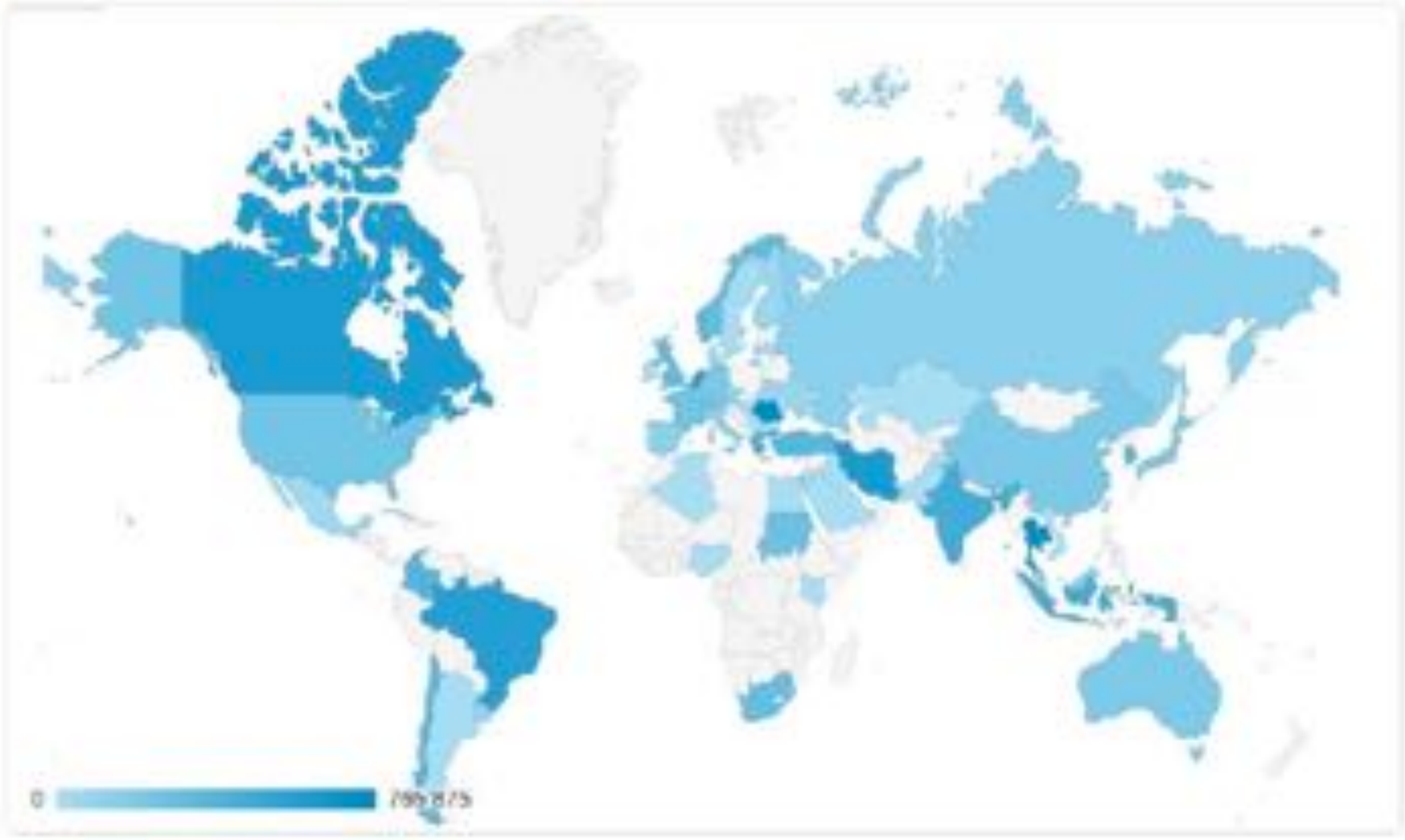
To get access to resources.

- compute power
- software
- storage
- data

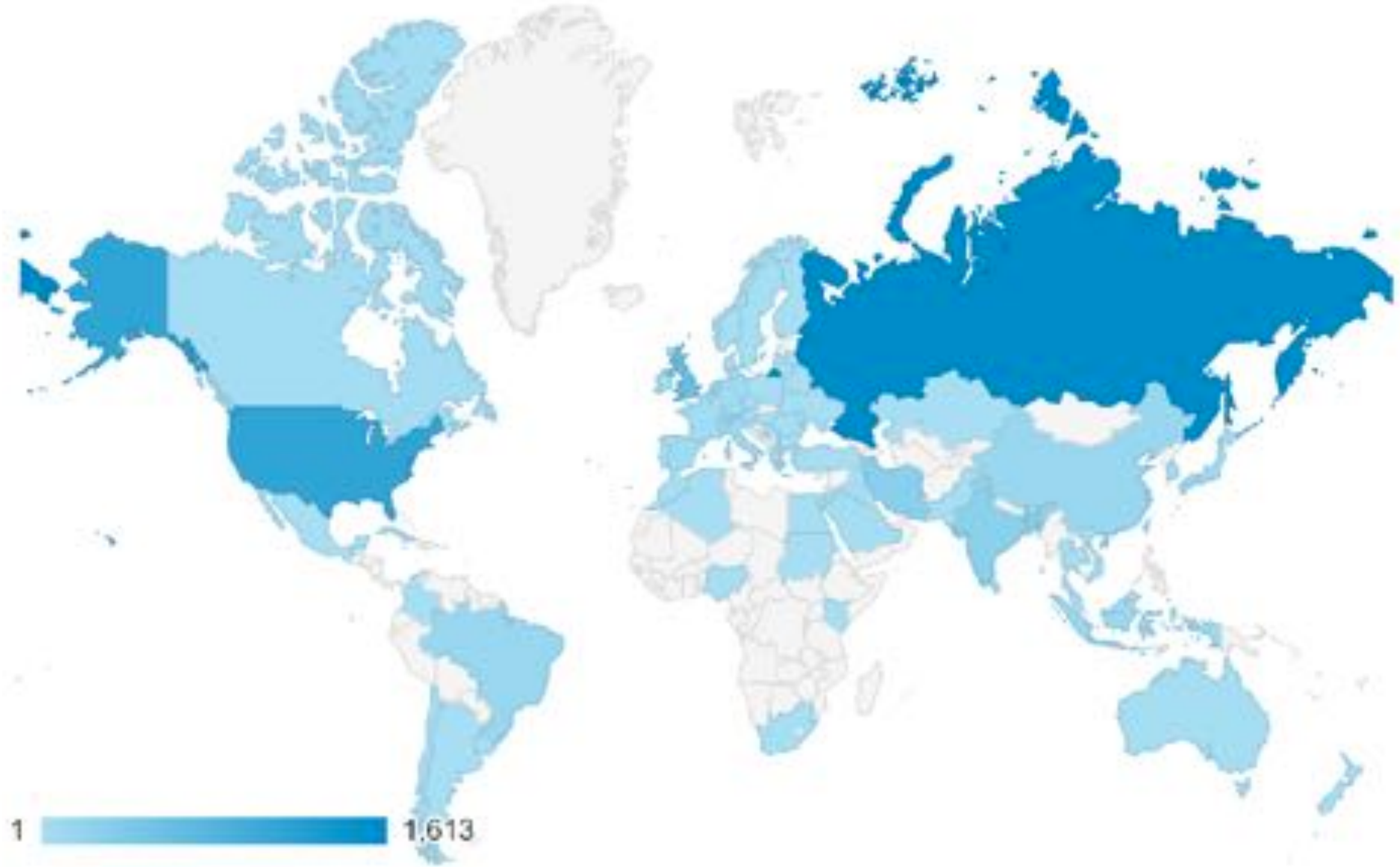
} *Academia is well-suited to serve as a micro-cloud provider of these resources*

David Koes @david_koes · 23 Sep 2016

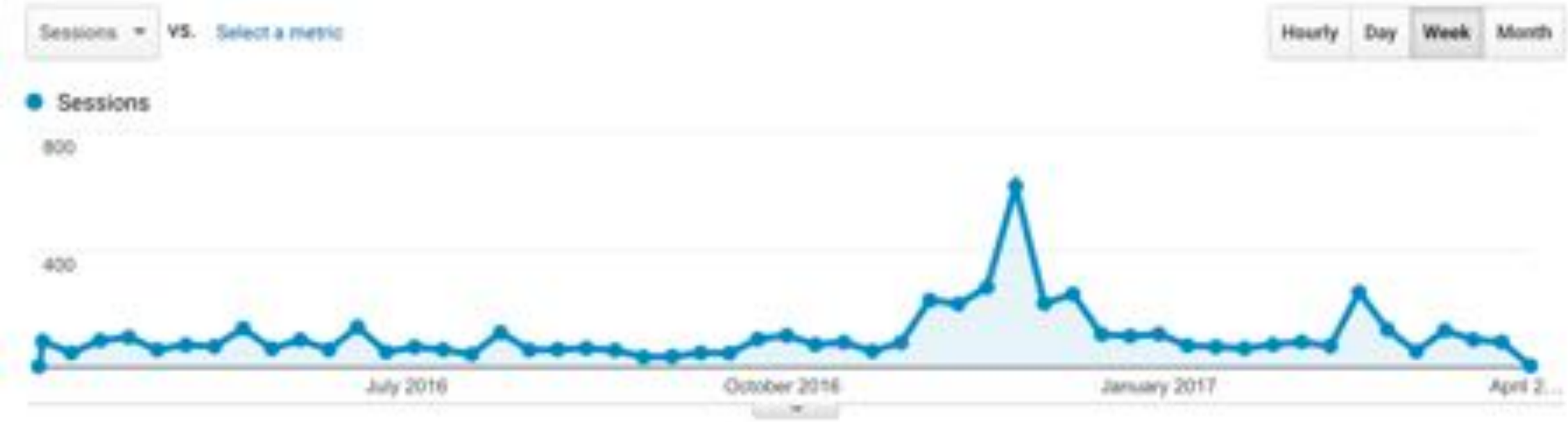
Pharmit usage by country. Online open drug discovery tools democratize access to big data resources.
pharmit.csb.pitt.edu #compchem



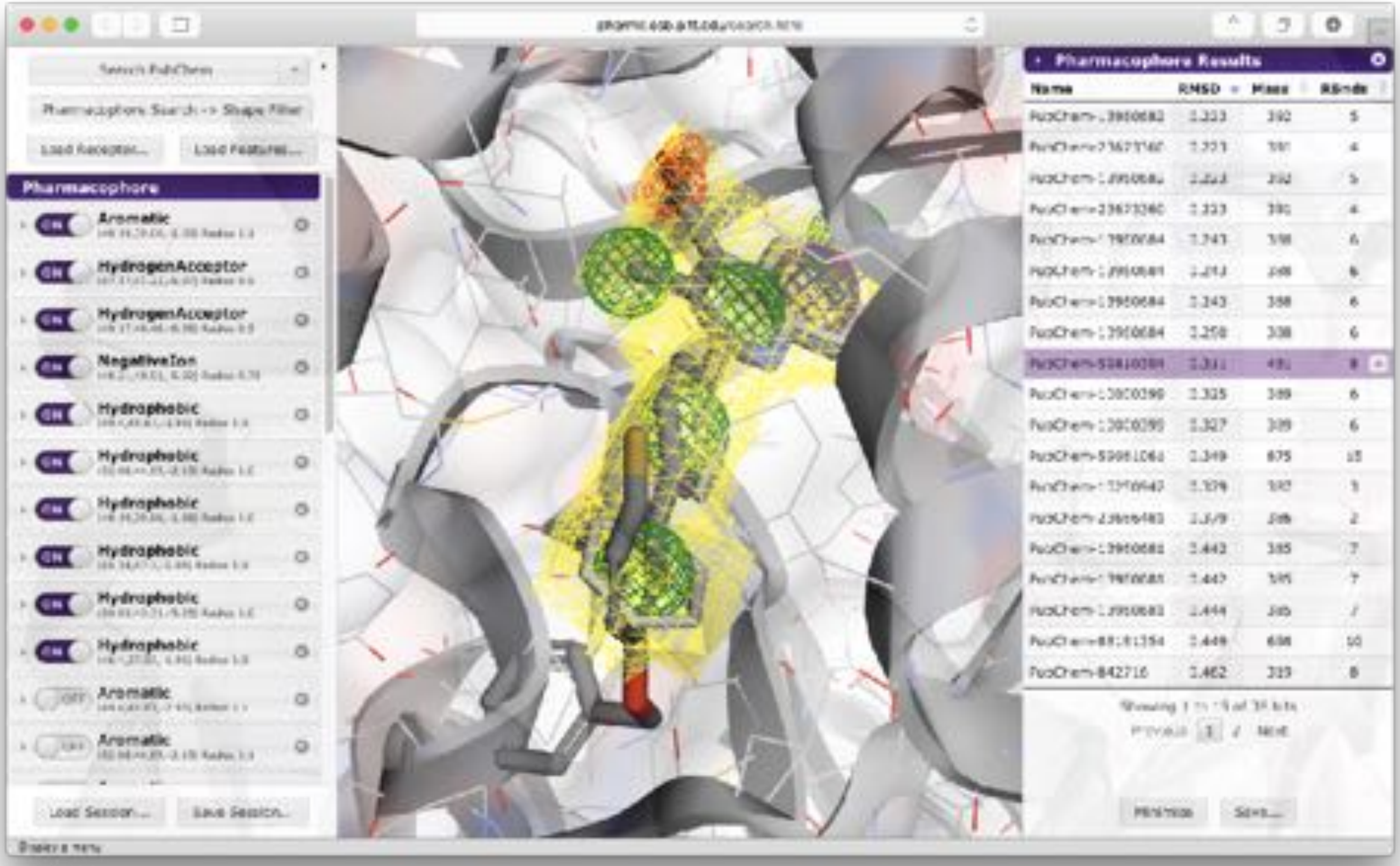
3 4


1
1,613

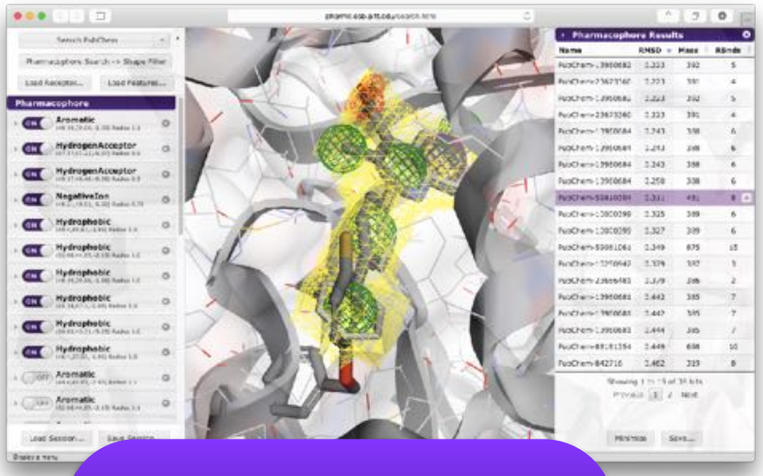
Country	Acquisition			Behavior		
	Sessions	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg. Session Duration
	5,604 % of Total: 100.00% (5,604)	52.57% Avg for View: 52.48% (0.17%)	2,946 % of Total: 100.17% (2,941)	28.09% Avg for View: 28.09% (0.00%)	3.08 Avg for View: 3.08 (0.00%)	00:05:07 Avg for View: 00:05:07 (0.00%)
1. Lithuania	1 (0.02%)	100.00%	1 (0.03%)	0.00%	7.00	00:23:01
2. Romania	8 (0.14%)	37.50%	3 (0.10%)	25.00%	4.75	00:12:46
3. Netherlands	156 (2.78%)	21.79%	34 (1.15%)	25.64%	7.53	00:12:43
4. Bangladesh	77 (1.37%)	28.57%	22 (0.75%)	37.66%	3.04	00:09:52
5. Iran	287 (5.12%)	31.71%	91 (3.09%)	23.00%	4.59	00:08:51
6. Indonesia	119 (2.12%)	15.97%	19 (0.64%)	20.17%	3.86	00:08:07
7. Brazil	158 (2.82%)	41.14%	65 (2.21%)	27.22%	3.85	00:07:48
8. United Kingdom	623 (11.12%)	46.55%	290 (9.84%)	36.12%	3.91	00:06:53
9. India	283 (5.05%)	54.42%	154 (5.23%)	41.34%	3.49	00:06:45
10. United States	1,116 (19.91%)	42.38%	473 (16.06%)	39.61%	3.20	00:06:08


Sessions
vs.
Select a metric
Hourly
Day
Week
Month
Sessions
800
400
July 2016
October 2016
January 2017
April 2...
Sessions
5,604
Users
2,966
Pageviews
17,274
Pages / Session
3.08
Avg. Session Duration
00:05:07
Bounce Rate
28.09%


New Visitor
Returning Visitor
47.4%
52.6%



http://pharmit.sf.net



pharmit.js

JQuery

3Dmol.js

Client

Search

User Libraries

CHEMBL

PubChem

ChemDiv

NCI

MolPort

Prebuilt Libraries

pharmitserver

OpenBabel

ShapeDB

Pharmer

Minimization

smina

OpenBabel

buildlibs.py

RDKit

OpenBabel

Server

http://pharmit.sf.net

GPL/BSD License

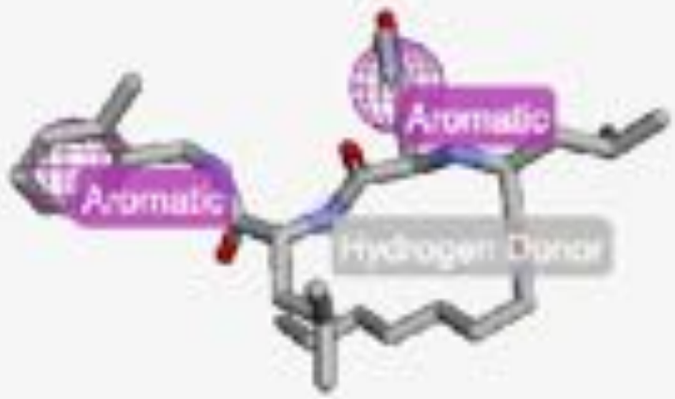
Client

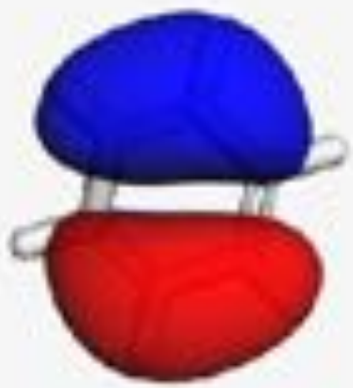
Why web application?

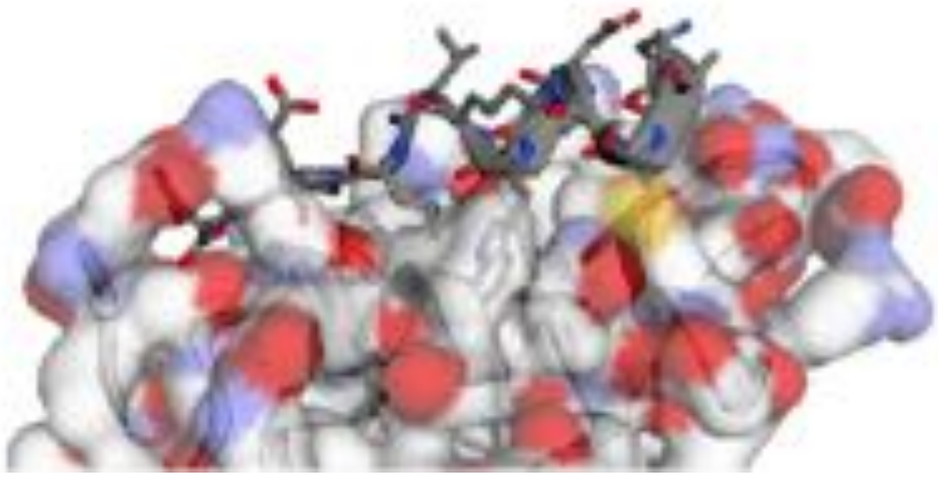
3Dmol.js

View Embed Develop Feedback Download

A modern, object-oriented JavaScript library for visualizing molecular data

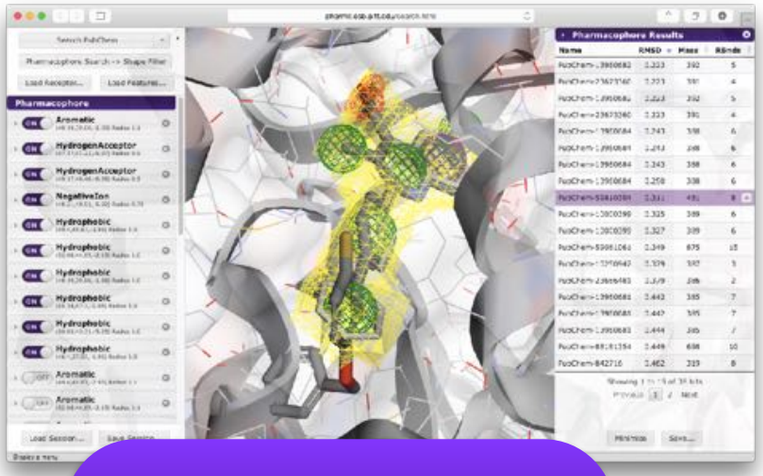






http://3dmol.csb.pitt.edu

Rego, N., & Koes, D. (2014). 3Dmol.js: molecular visualization with WebGL. *Bioinformatics*. doi:10.1093/bioinformatics/btu829



pharmit.js

JQuery

3Dmol.js

Client

Search

User Libraries

CHEMBL

PubChem

ChemDiv

NCI

MolPort

Prebuilt Libraries

pharmitserver

OpenBabel

ShapeDB

Pharmer

Minimization

smina

OpenBabel

buildlibs.py

RDKit

OpenBabel

Server

http://pharmit.sf.net

CHEMBL

ChemDiv

MolPort

NCI

PubChem

CHEMBL22.1
20,828,750 conformers of 1,581,180 molecules
Updated: 2017-Feb-24 12:37:33

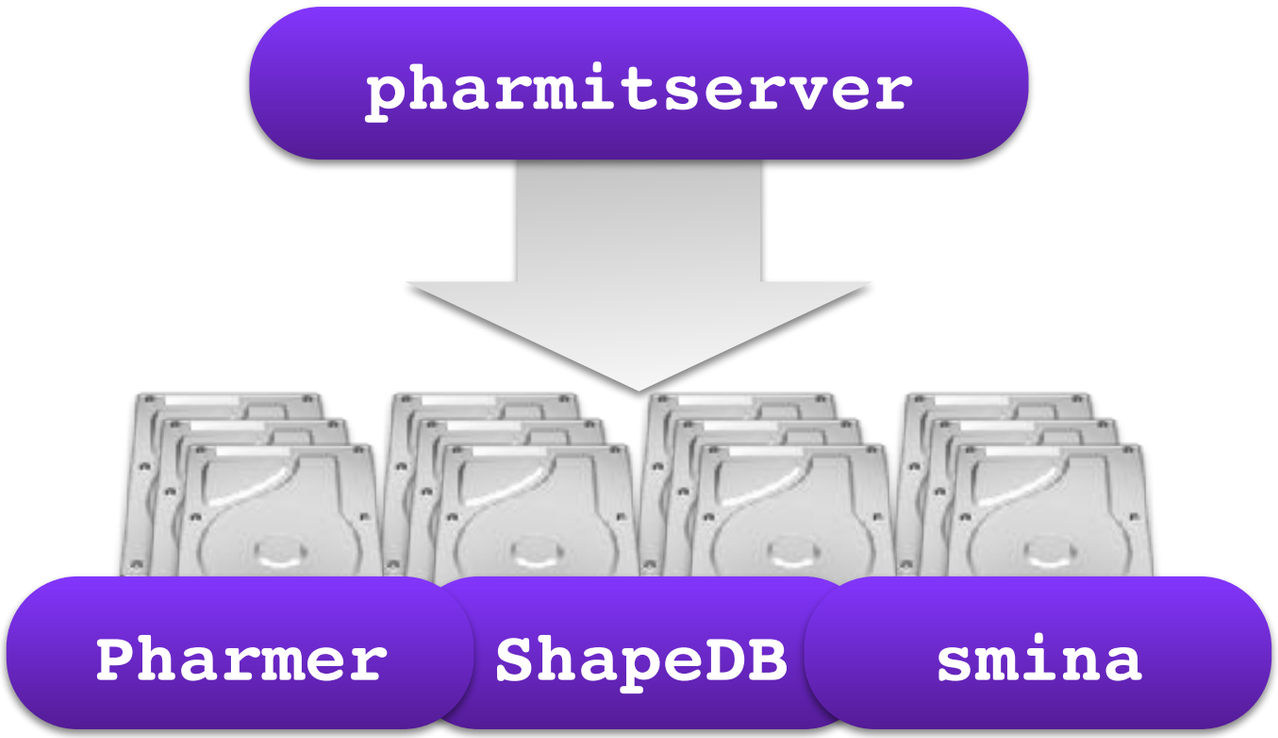
ChemDiv
21,609,296 conformers of 1,456,338 molecules
Updated: 2017-Feb-23 23:25:40

MolPort
94,527,744 conformers of 6,544,814 molecules
Updated: 2017-Feb-26 19:55:13

NCI Open Chemical Repository
1,590,246 conformers of 150,388 molecules
Updated: 2017-Feb-23 10:37:26

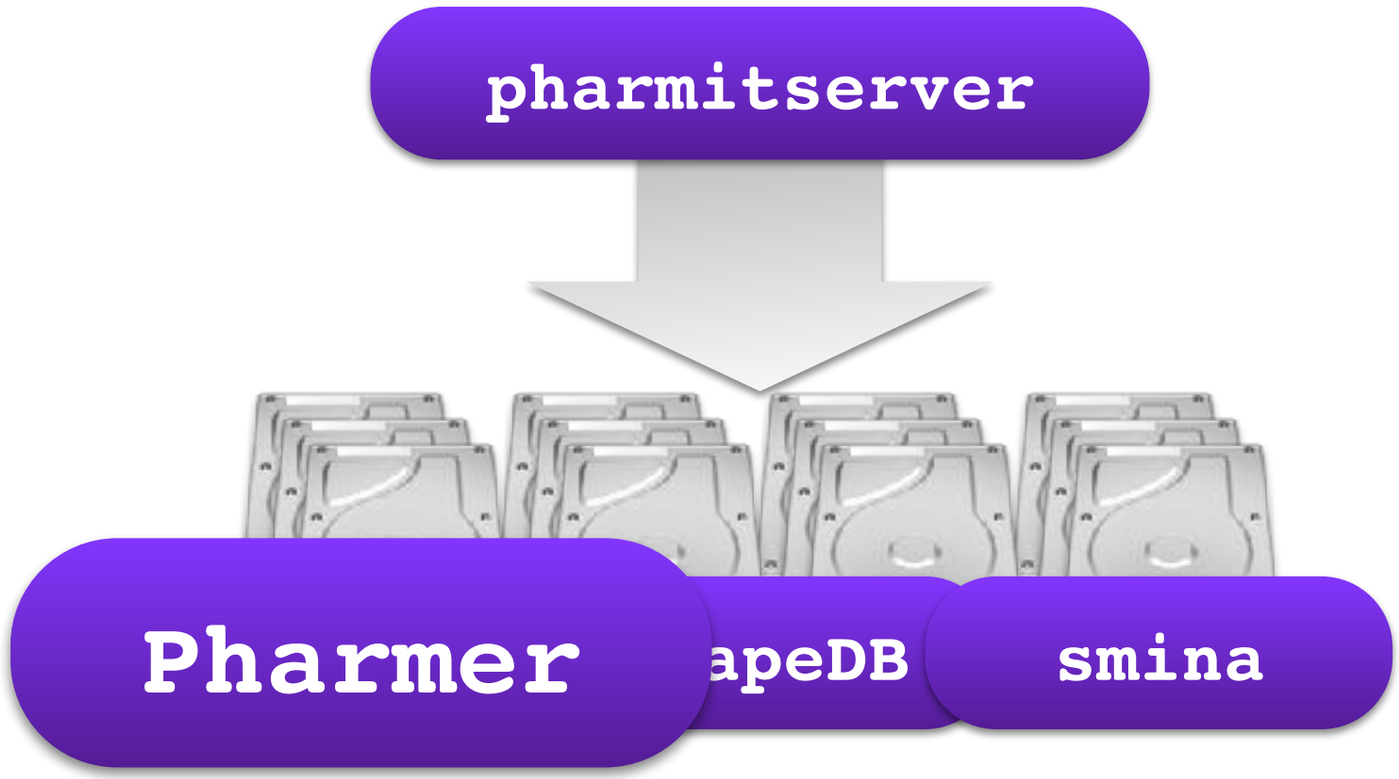
Contributed Libraries ›

Access Private Library



http://pharmit.sf.net

GPL/BSD License

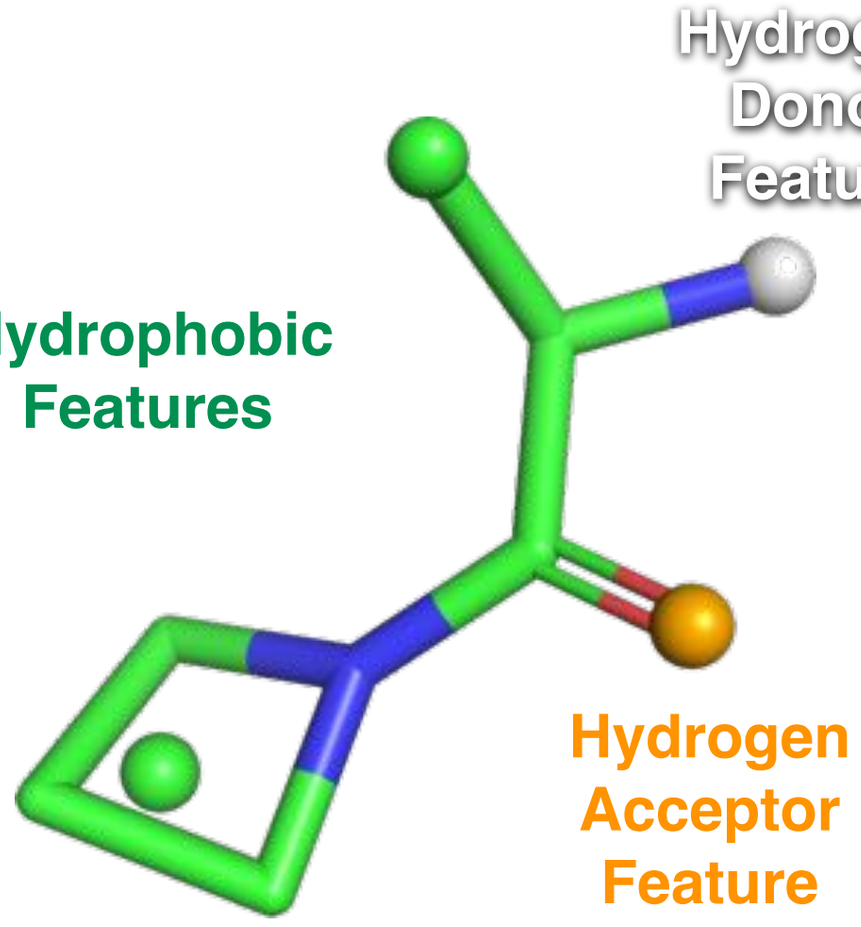


http://pharmit.sf.net

GPL/BSD License

Pharmer

Efficient and Exact Pharmacophore Search



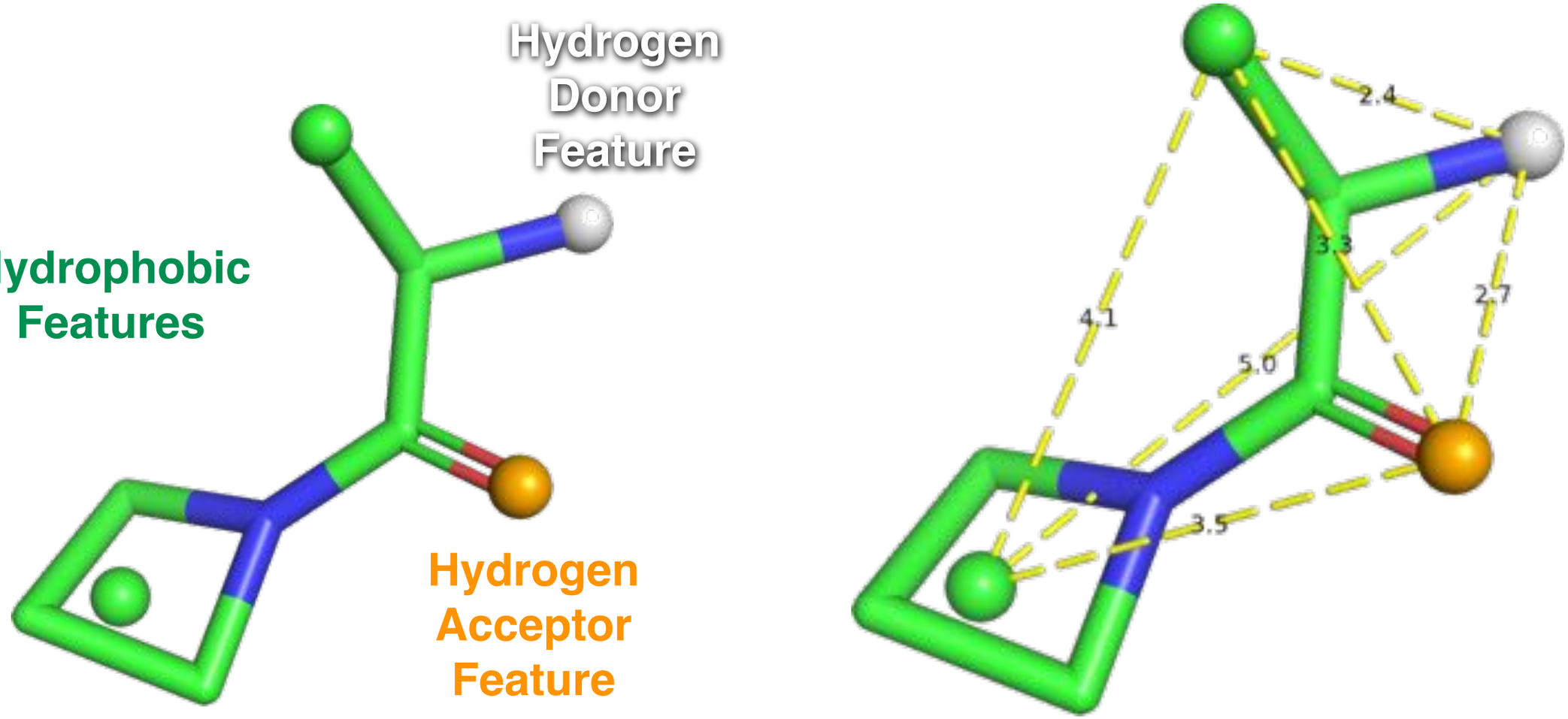
Pharmacophore

A spatial arrangement of molecular features essential for biological activity

Koes, D. R., & Camacho, C. J. (2011). Pharmer: efficient and exact pharmacophore search. *Journal of Chemical Information and Modeling, 51*(6), 1307-1314. doi:10.1021/ci200097m
Koes, D. R., & Camacho, C. J. (2012). ZINCPharmer: pharmacophore search of the ZINC database. *Nucleic acids research, 40*(Web Server issue), W409-414. doi:10.1093/nar/gks378

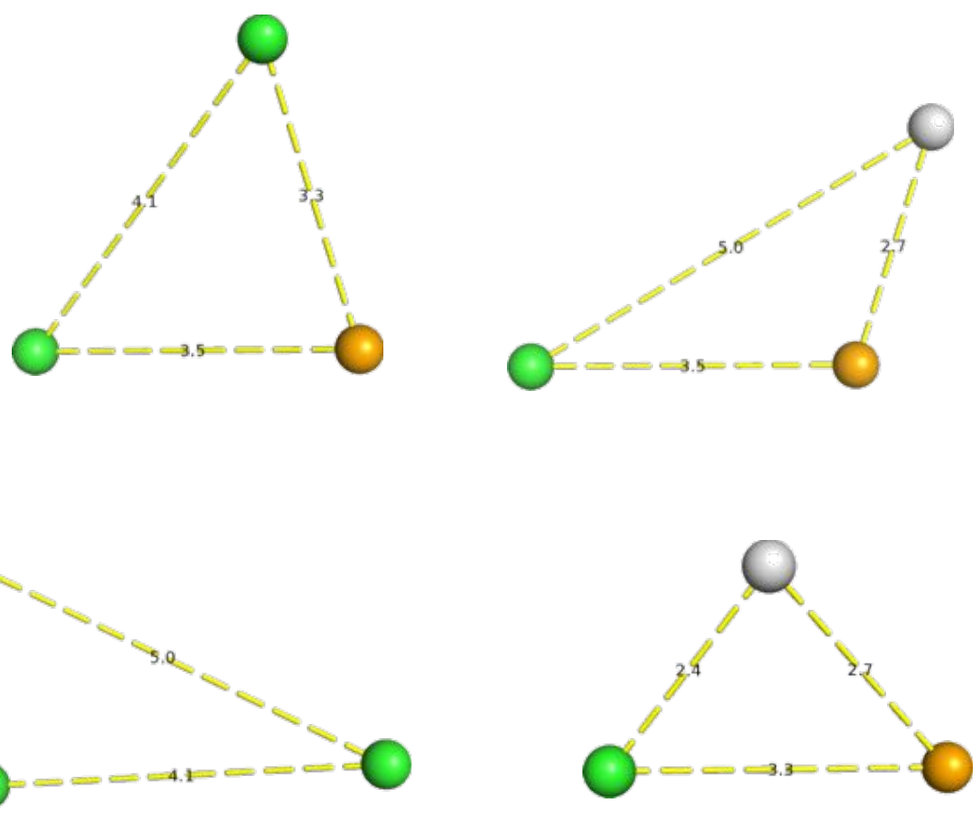
Pharmer

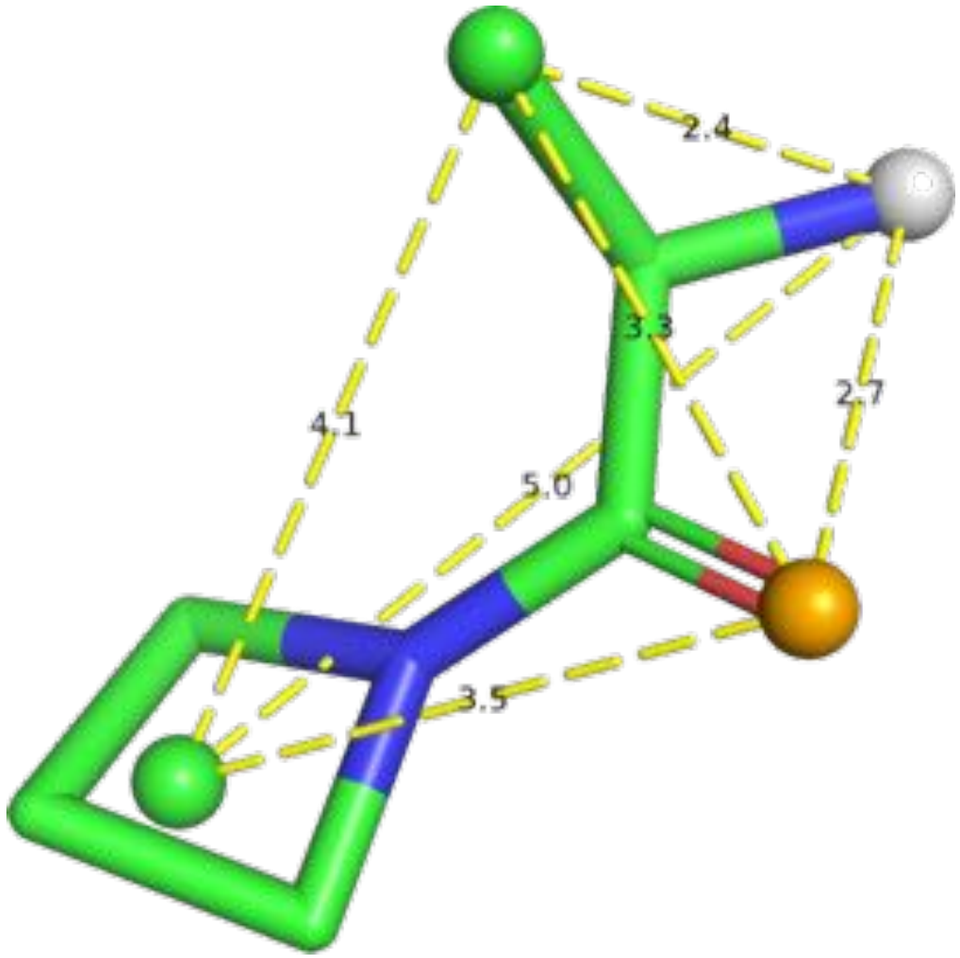
Efficient and Exact Pharmacophore Search



Pharmer

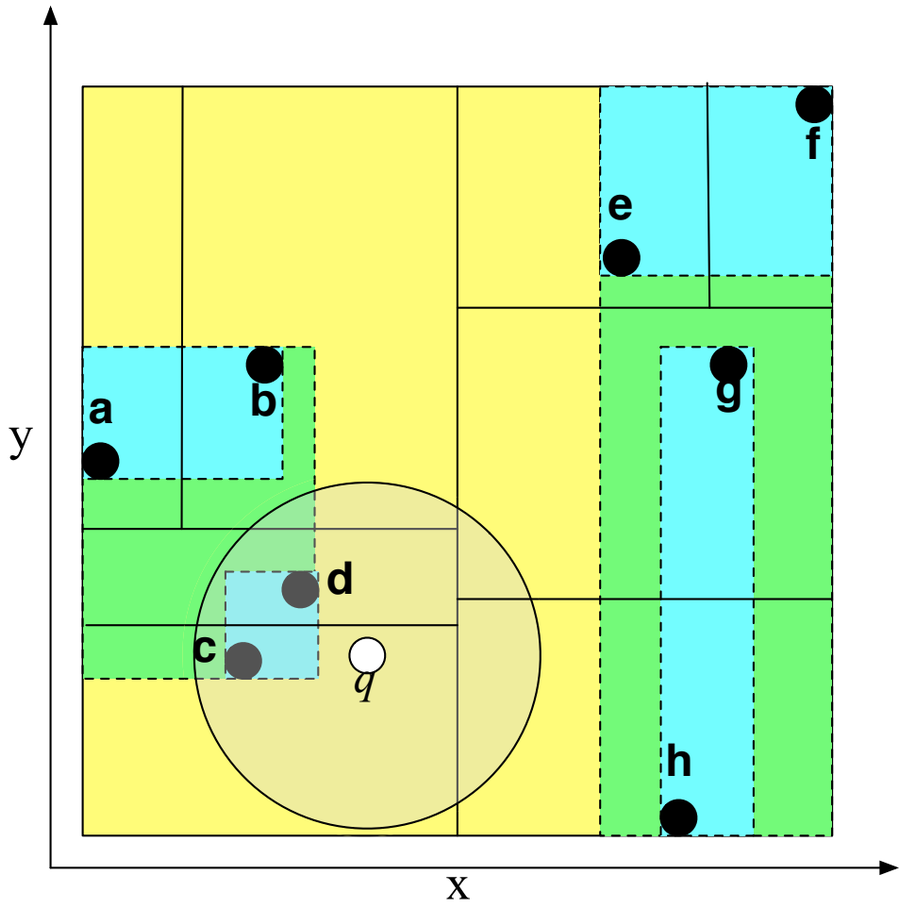
Efficient and Exact Pharmacophore Search

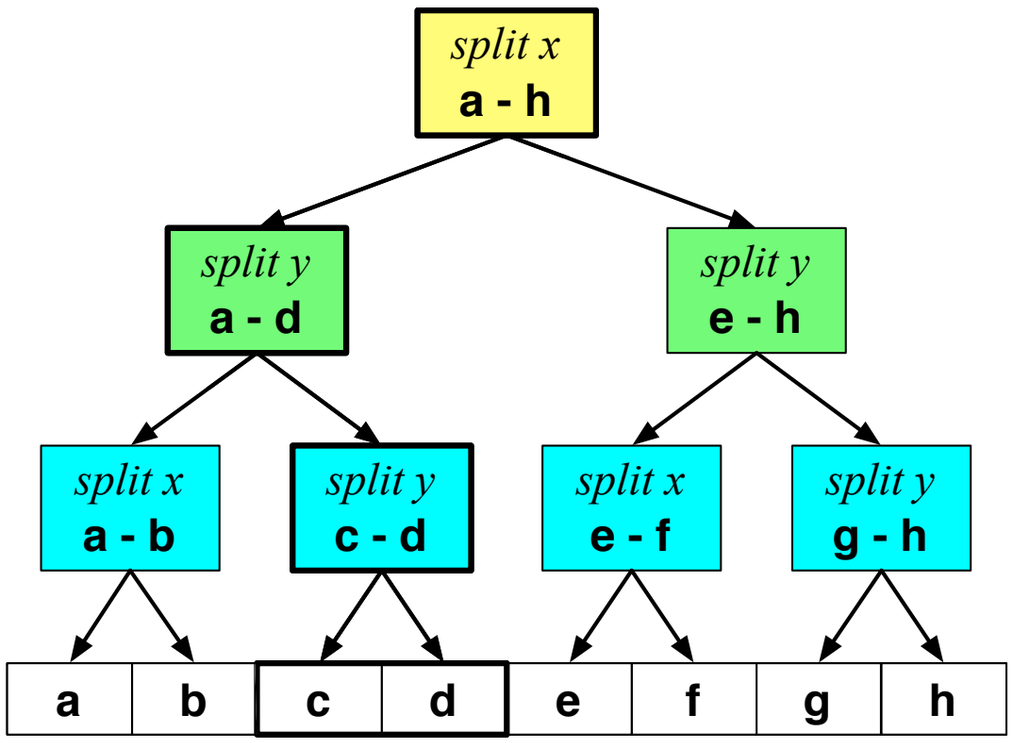




Pharmer

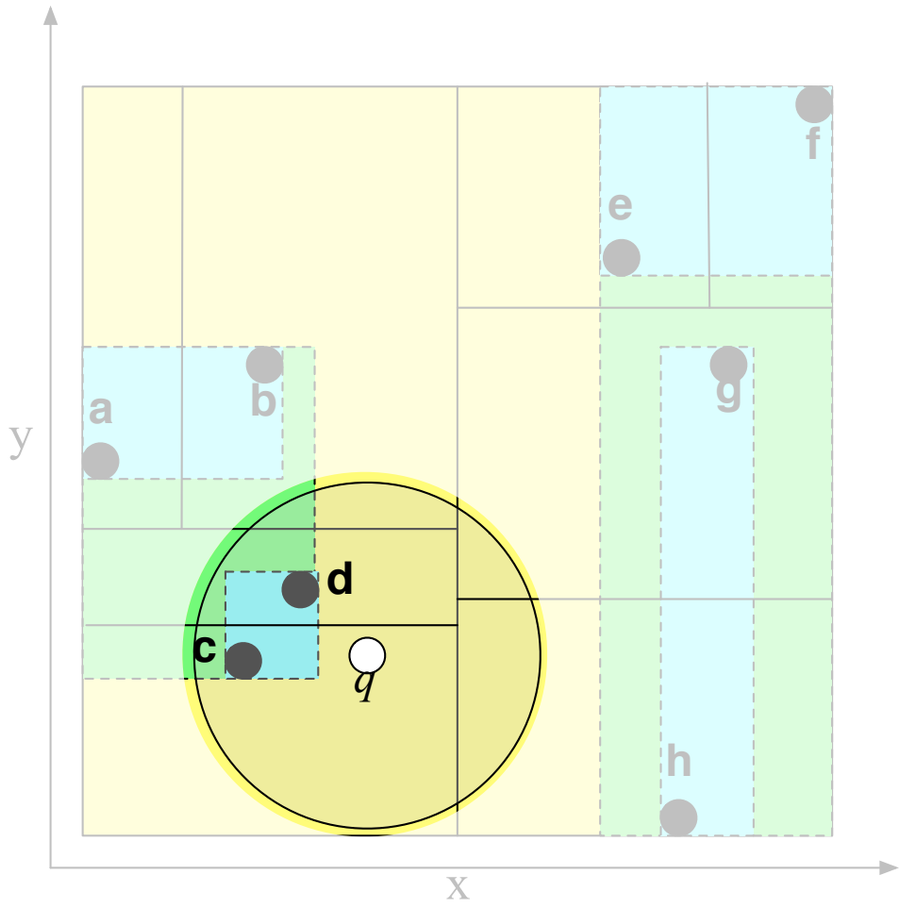
Efficient and Exact Pharmacophore Search

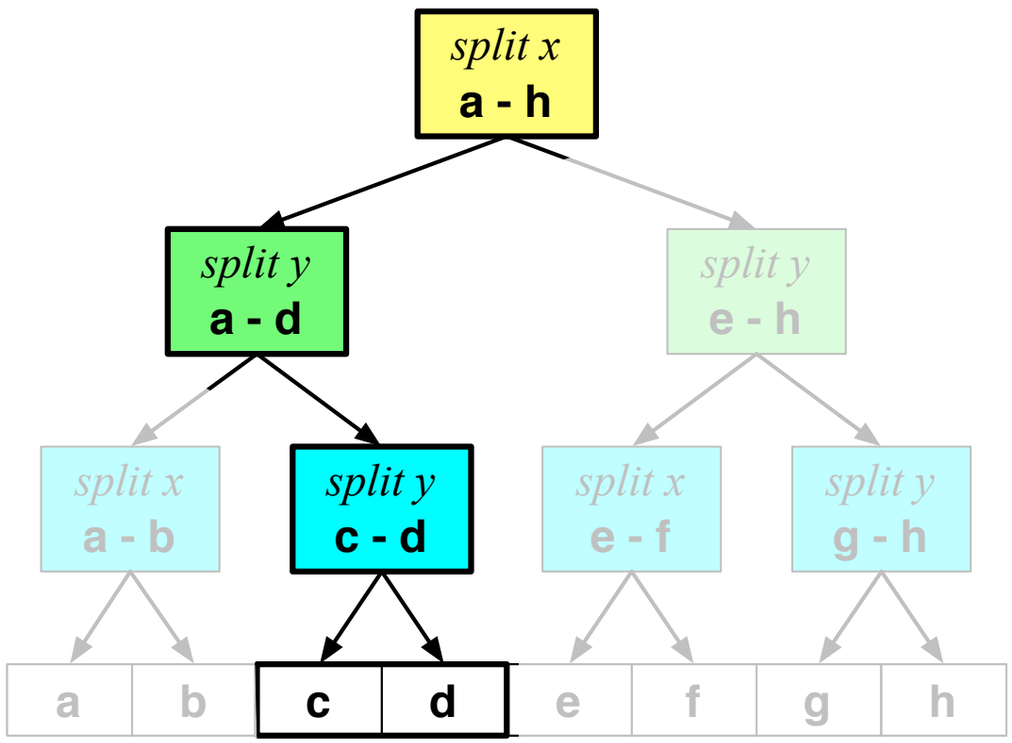




Pharmer

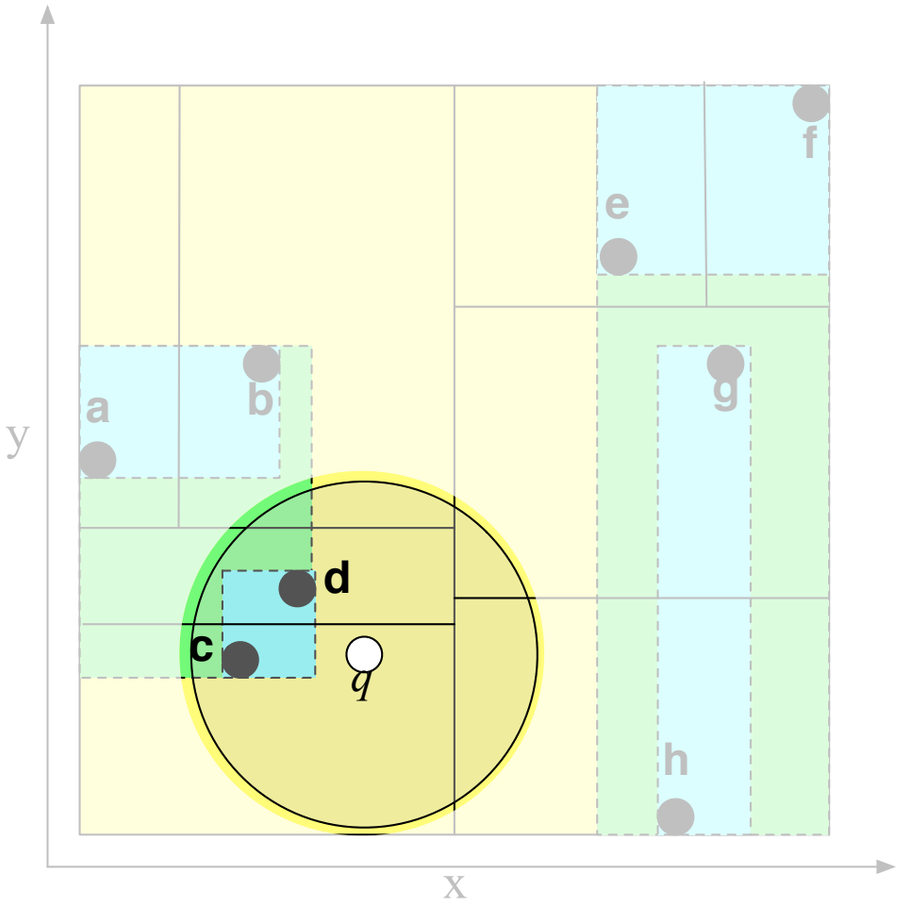
Efficient and Exact Pharmacophore Search

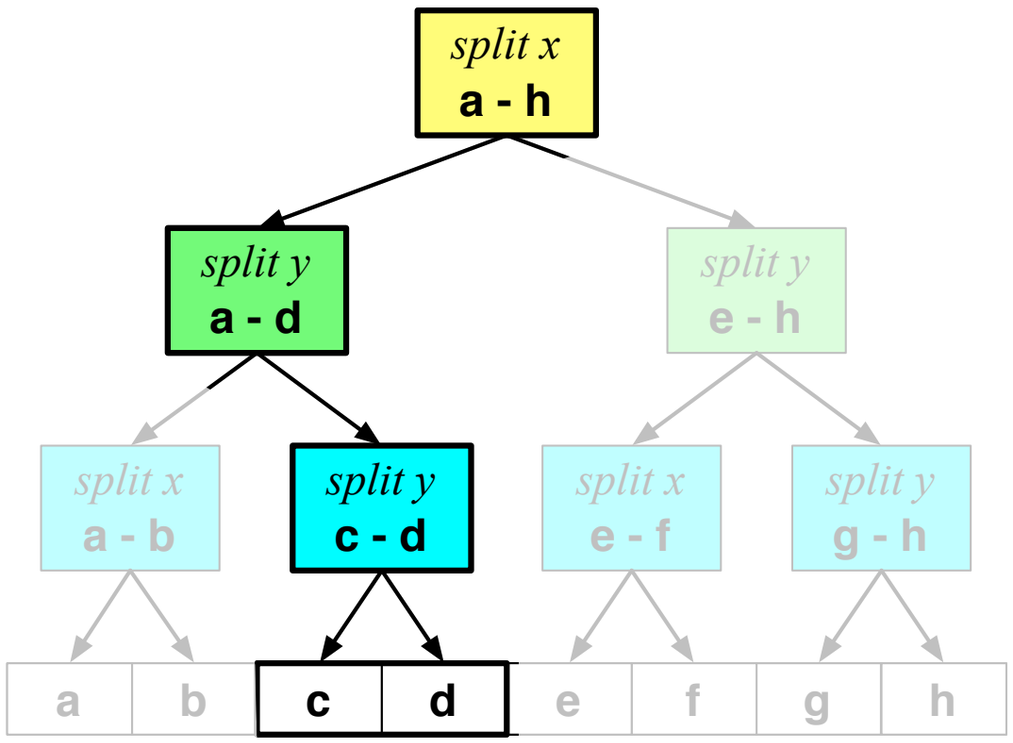


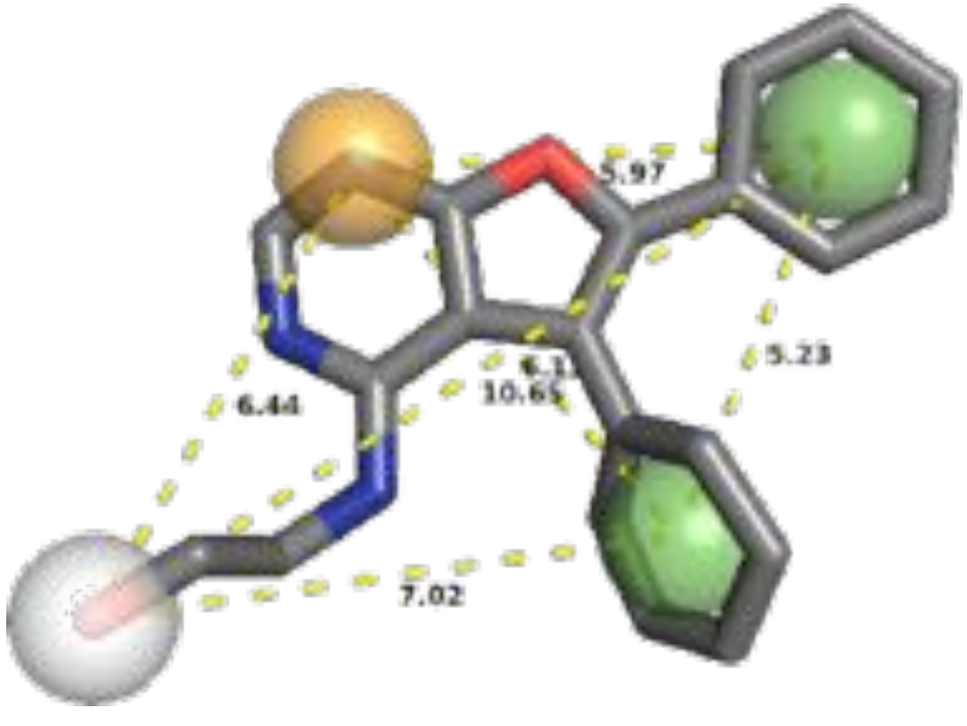


Pharmer

Efficient and Exact Pharmacophore Search

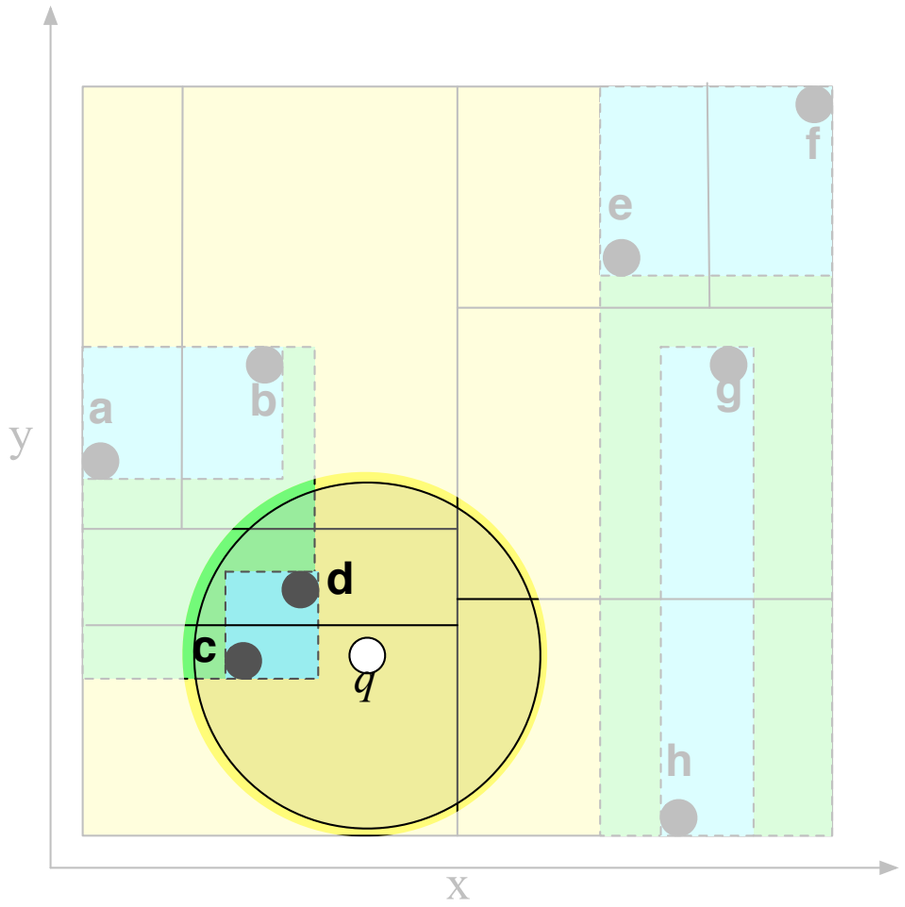


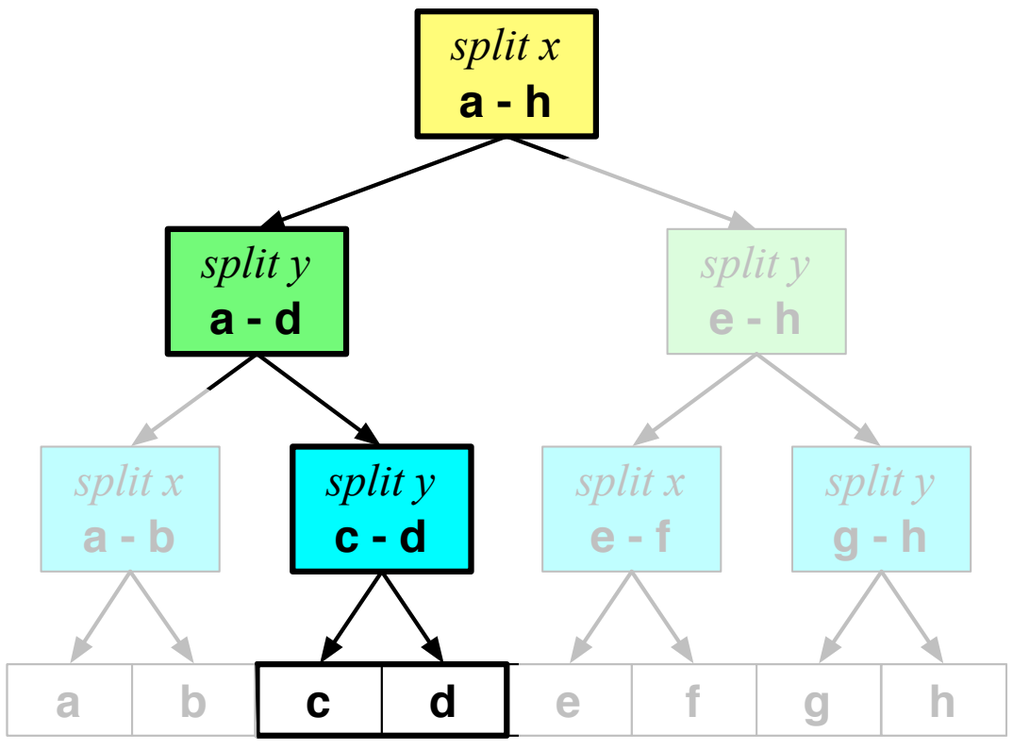


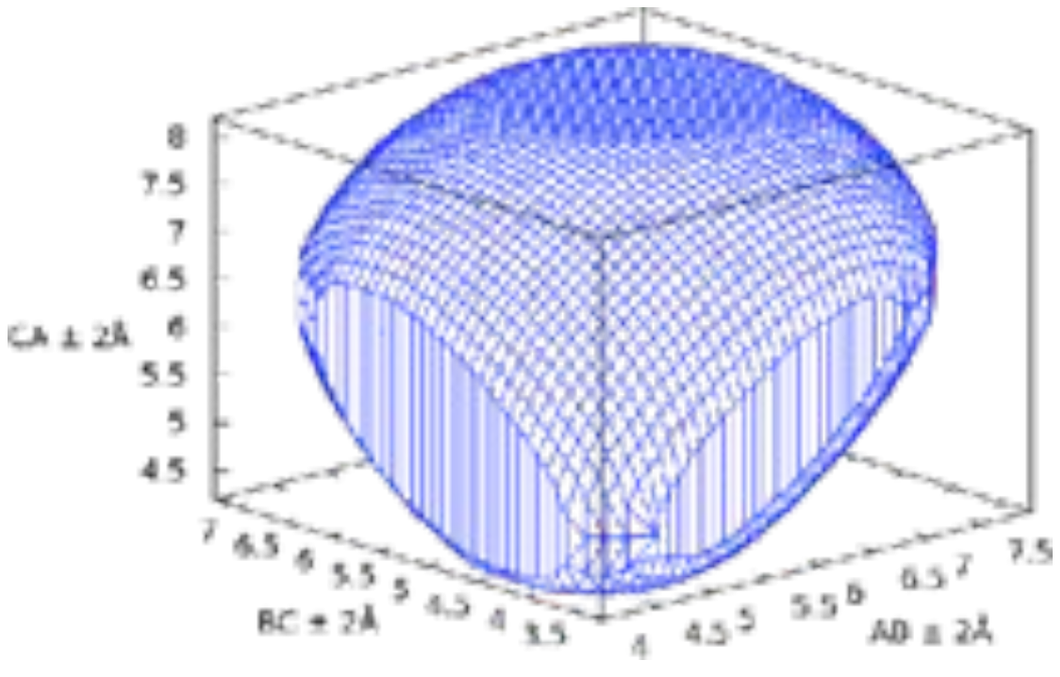


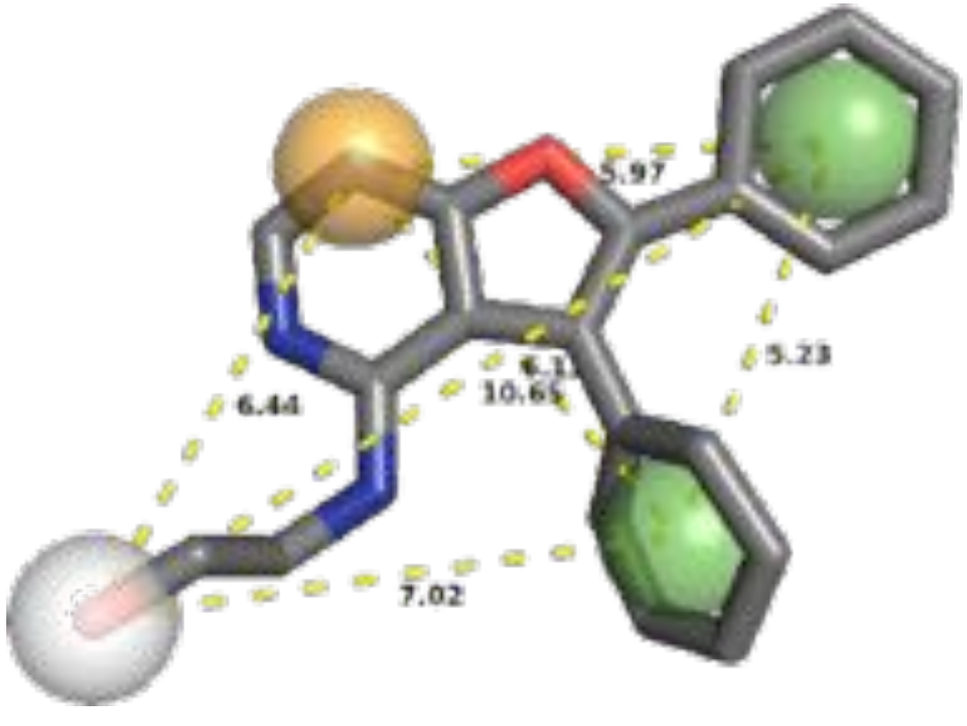
Pharmer

Efficient and Exact Pharmacophore Search



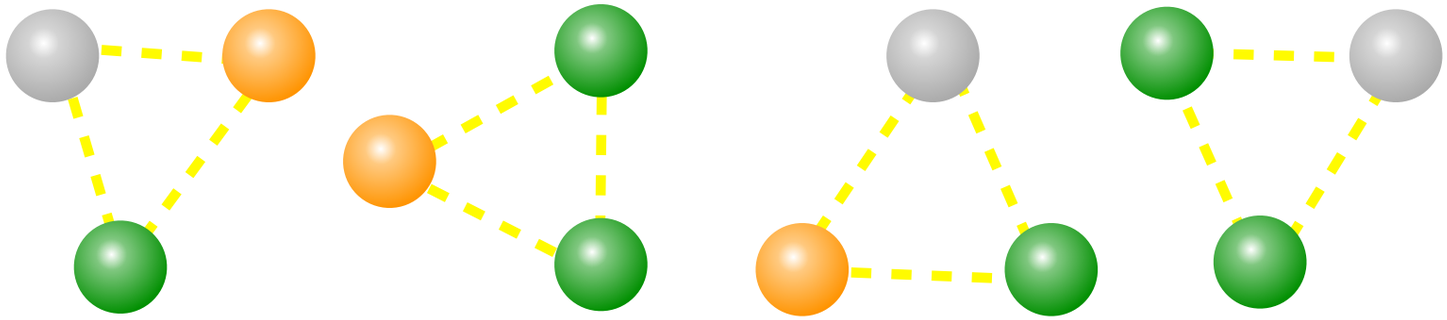


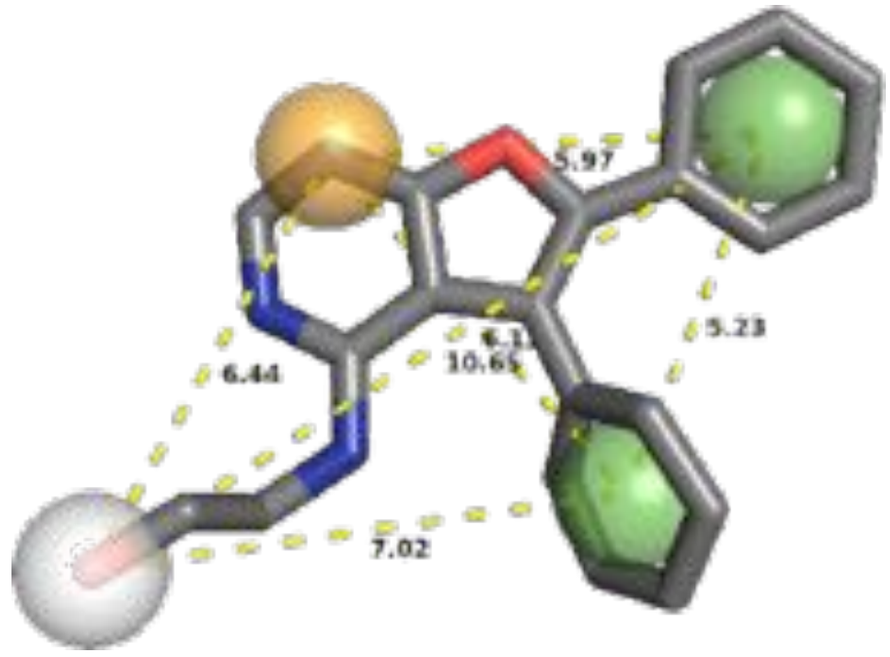




Pharmer

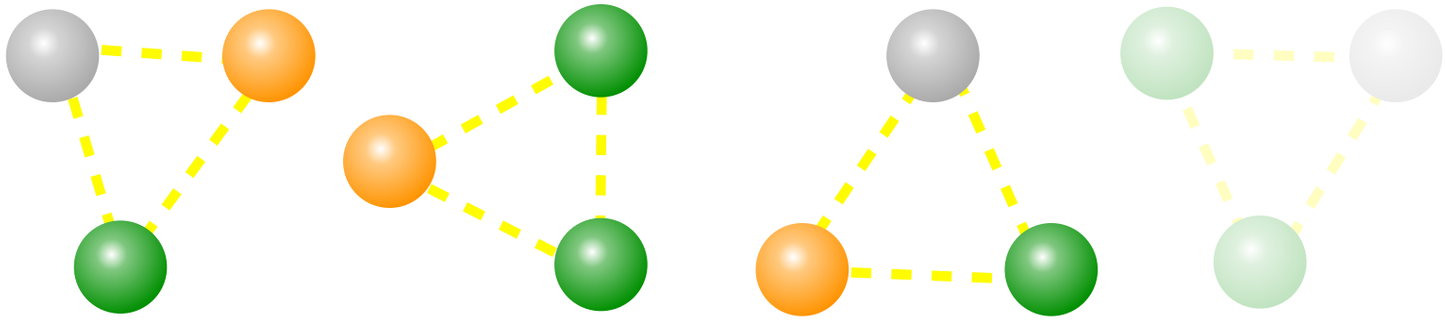
Efficient and Exact Pharmacophore Search

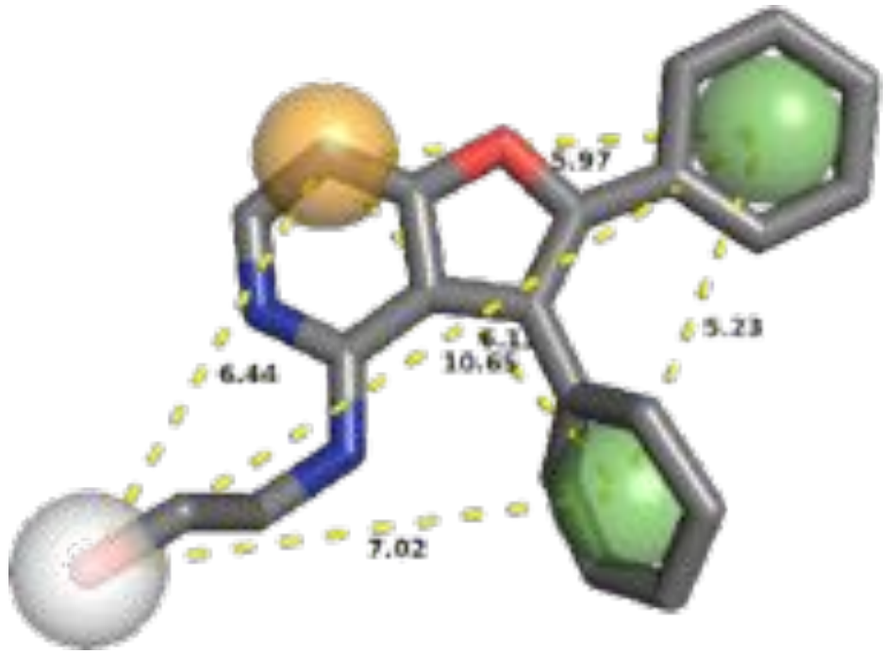




Pharmer

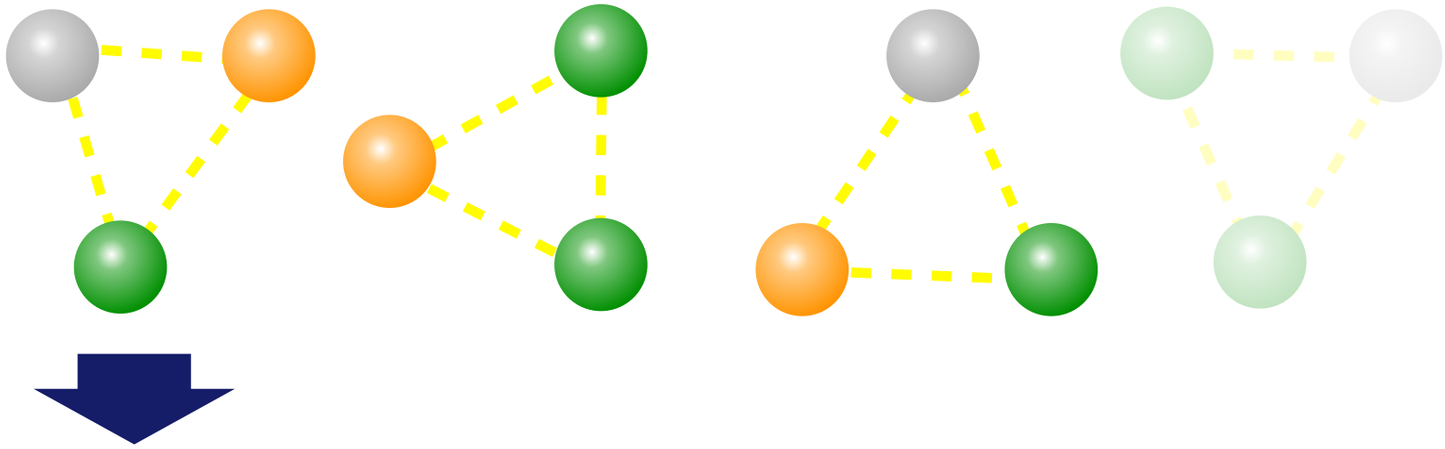
Efficient and Exact Pharmacophore Search

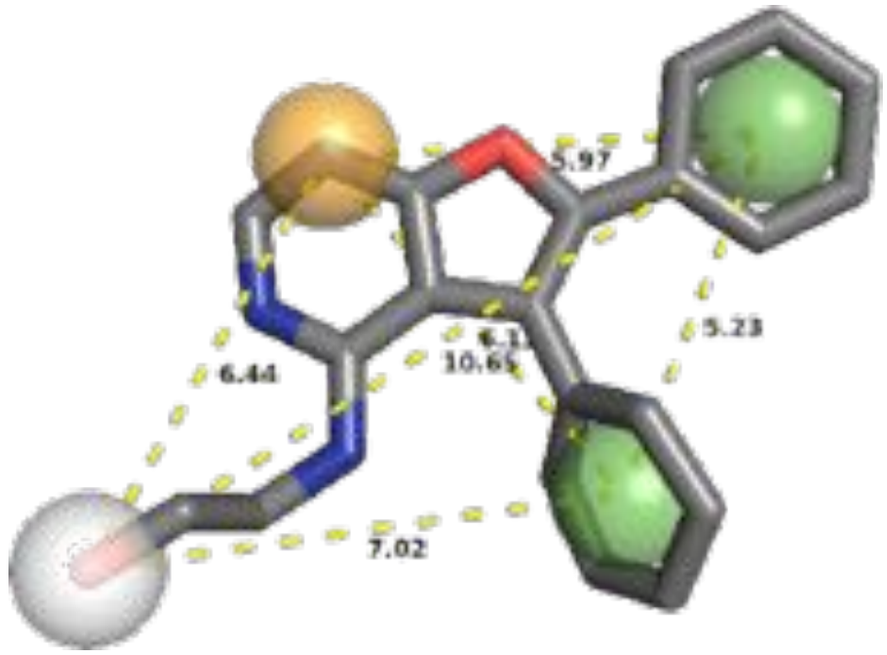




Pharmer

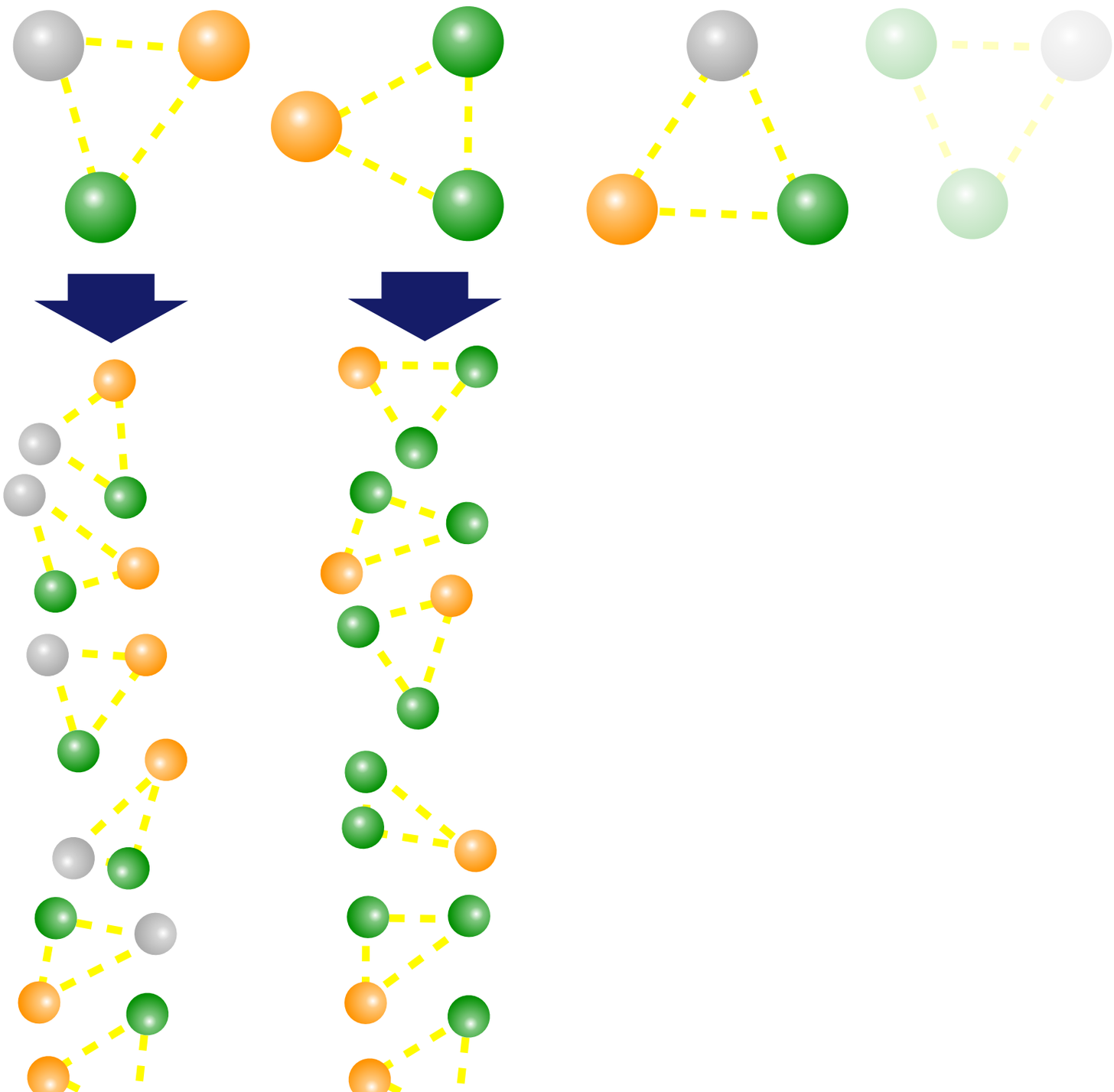
Efficient and Exact Pharmacophore Search

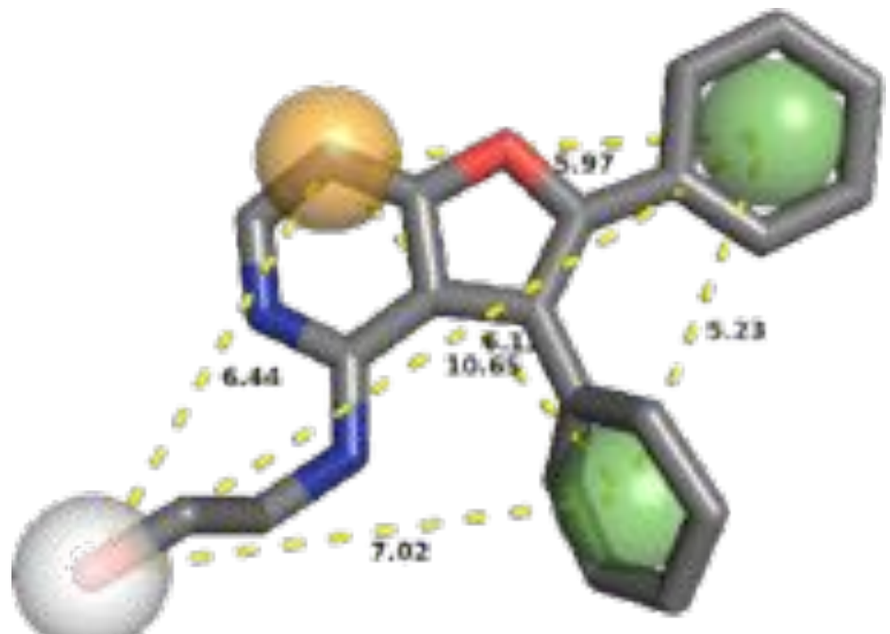




Pharmer

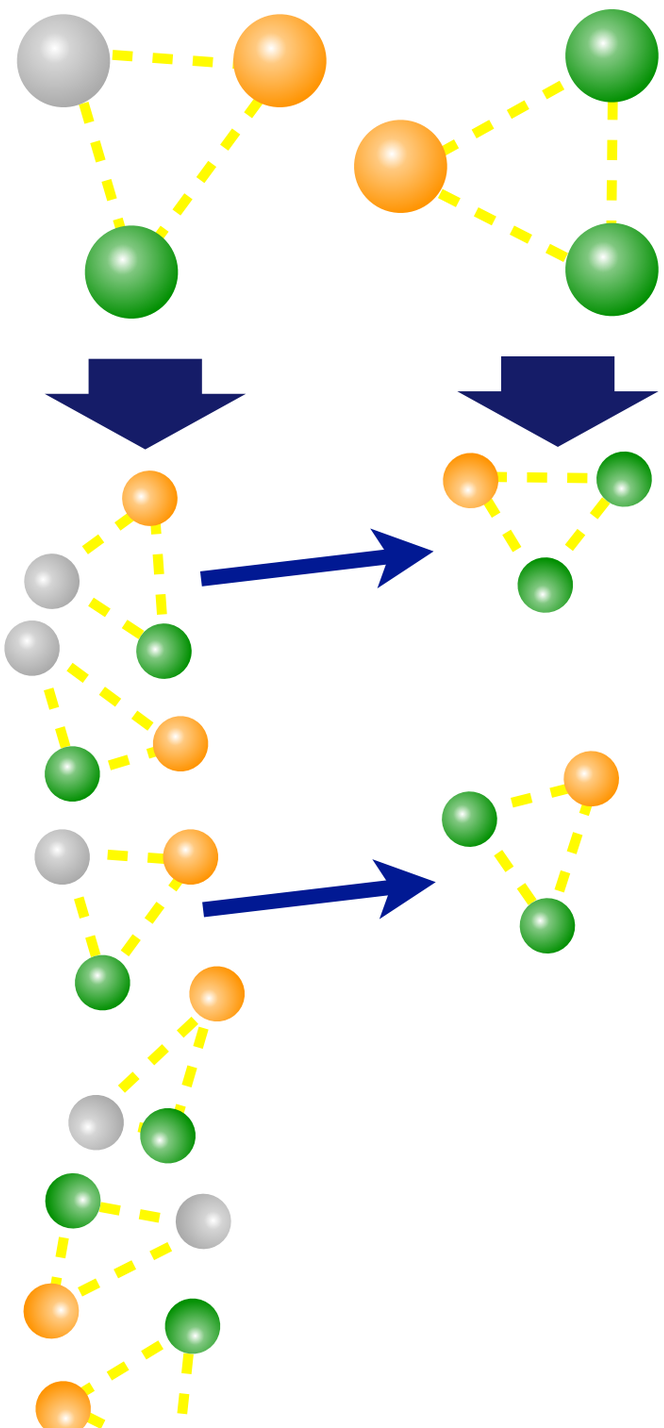
Efficient and Exact Pharmacophore Search

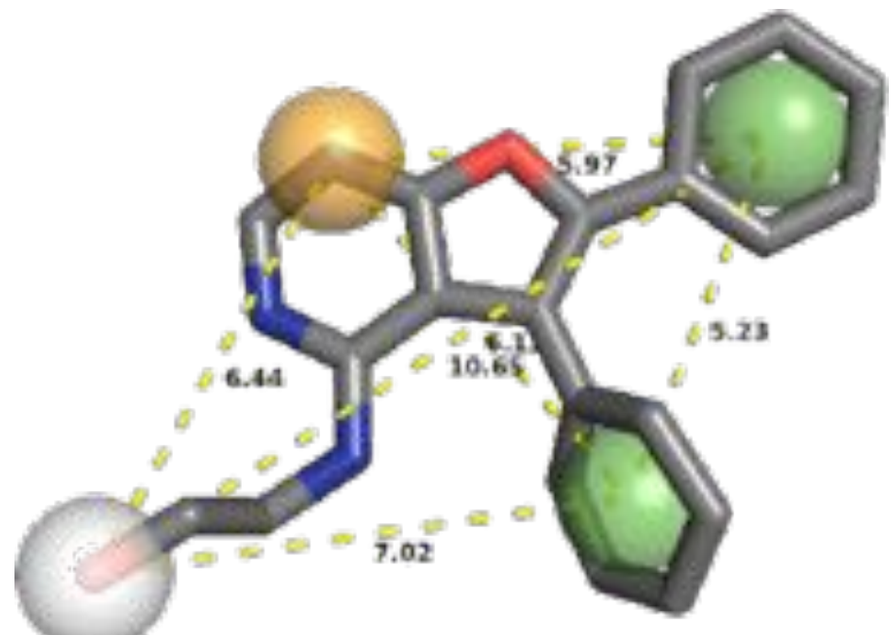




Pharmer

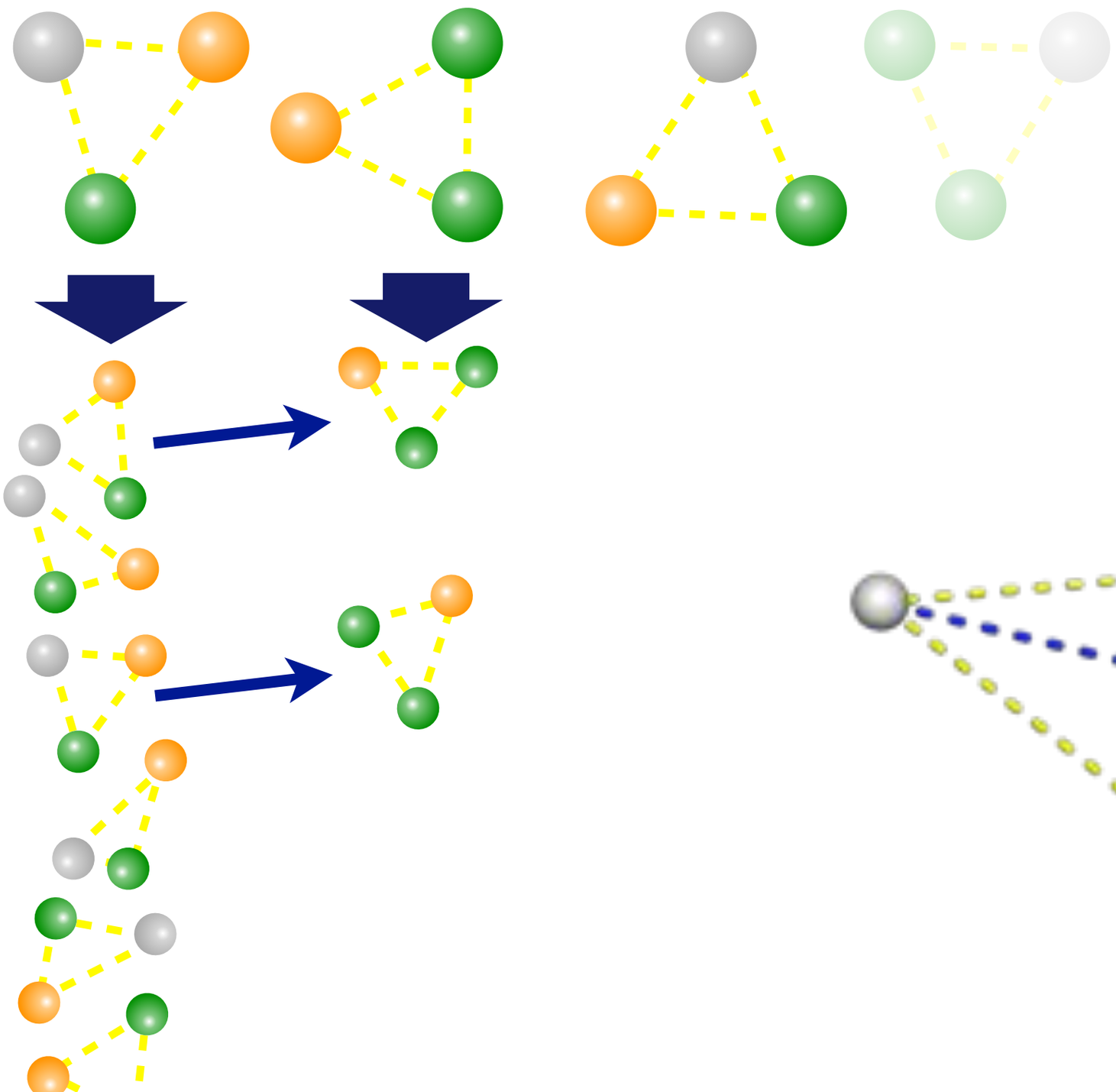
Efficient and Exact Pharmacophore Search

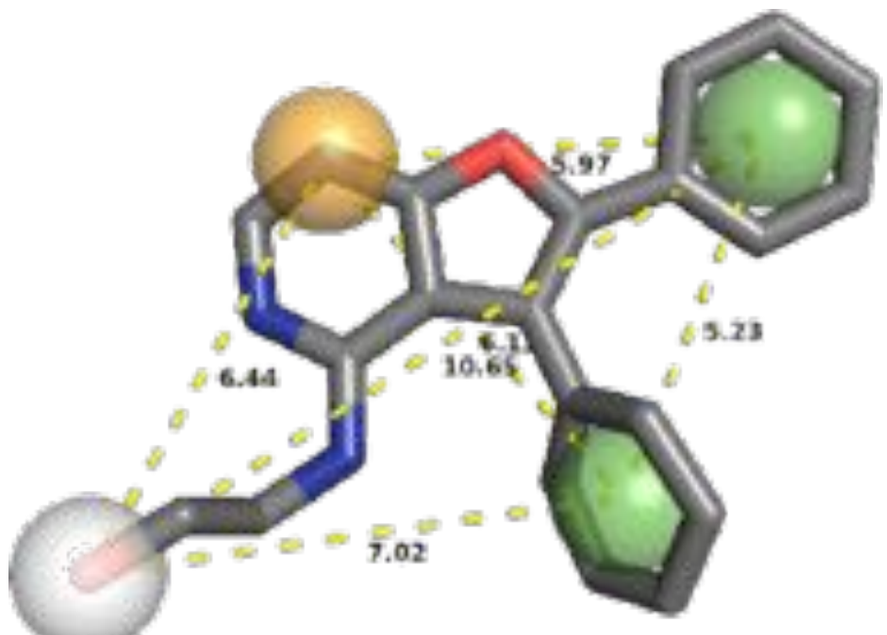


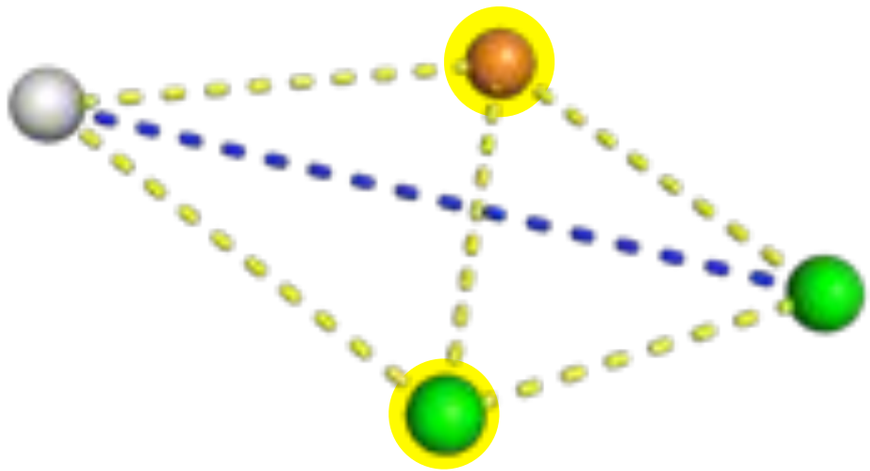


Pharmer

Efficient and Exact Pharmacophore Search

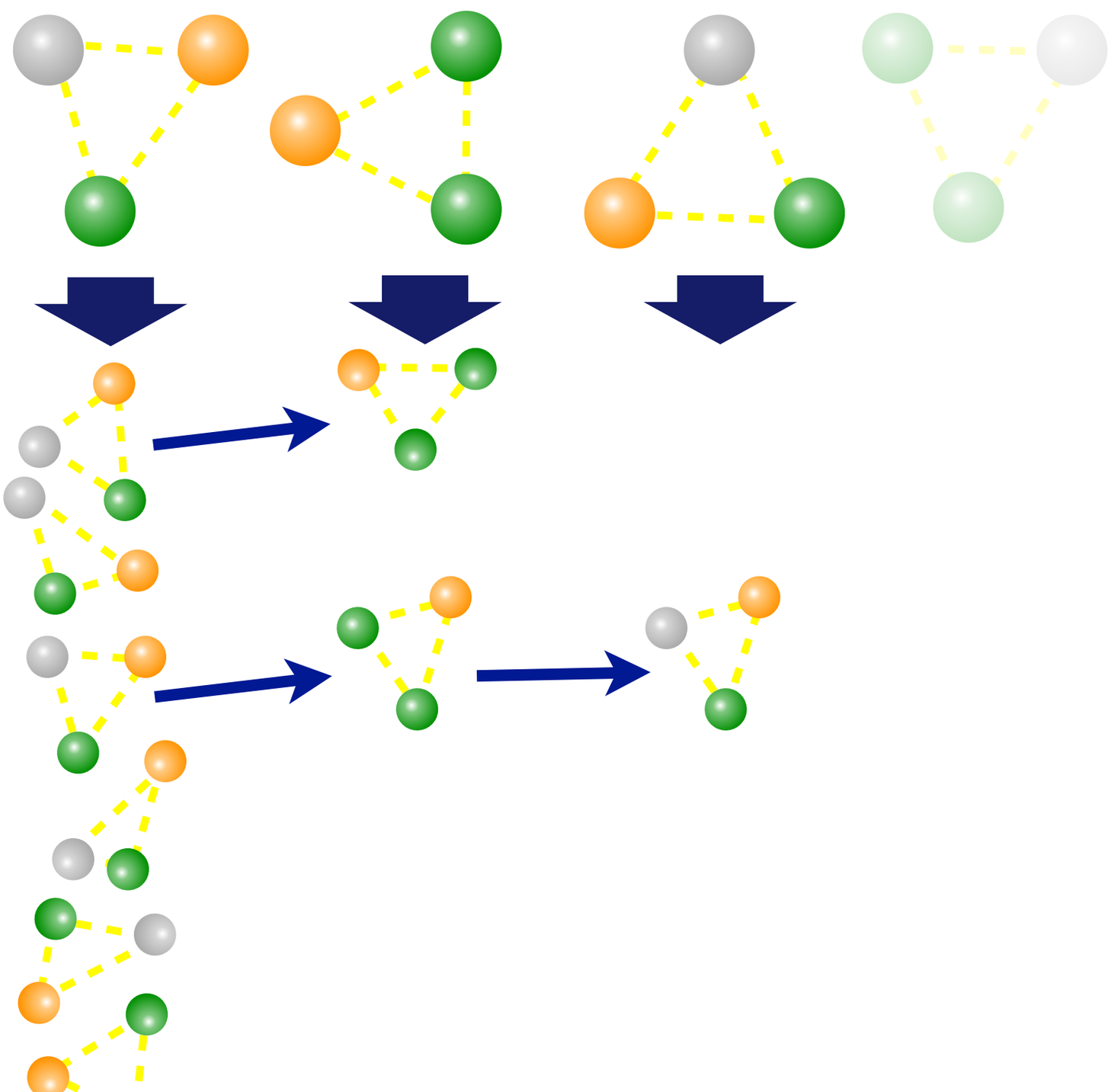


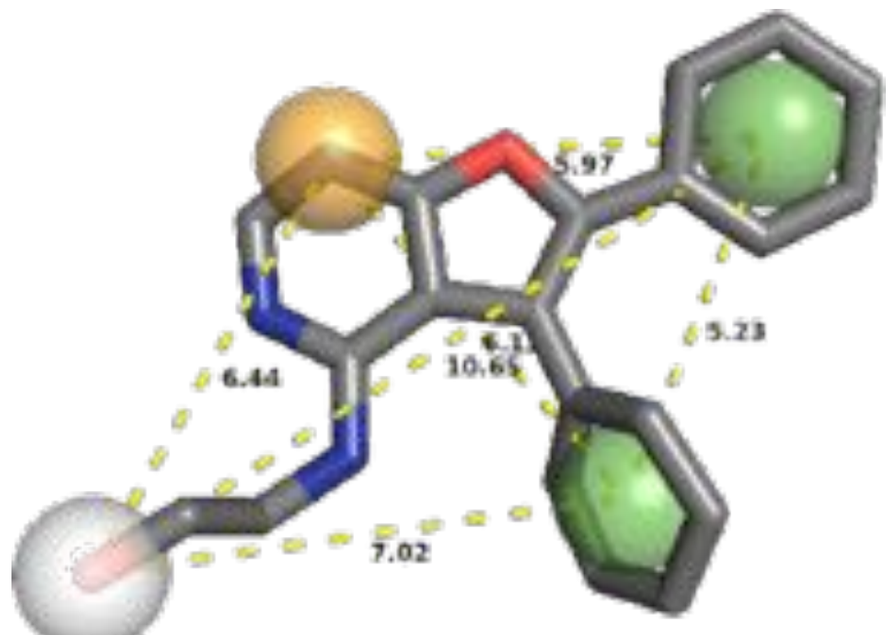




Pharmer

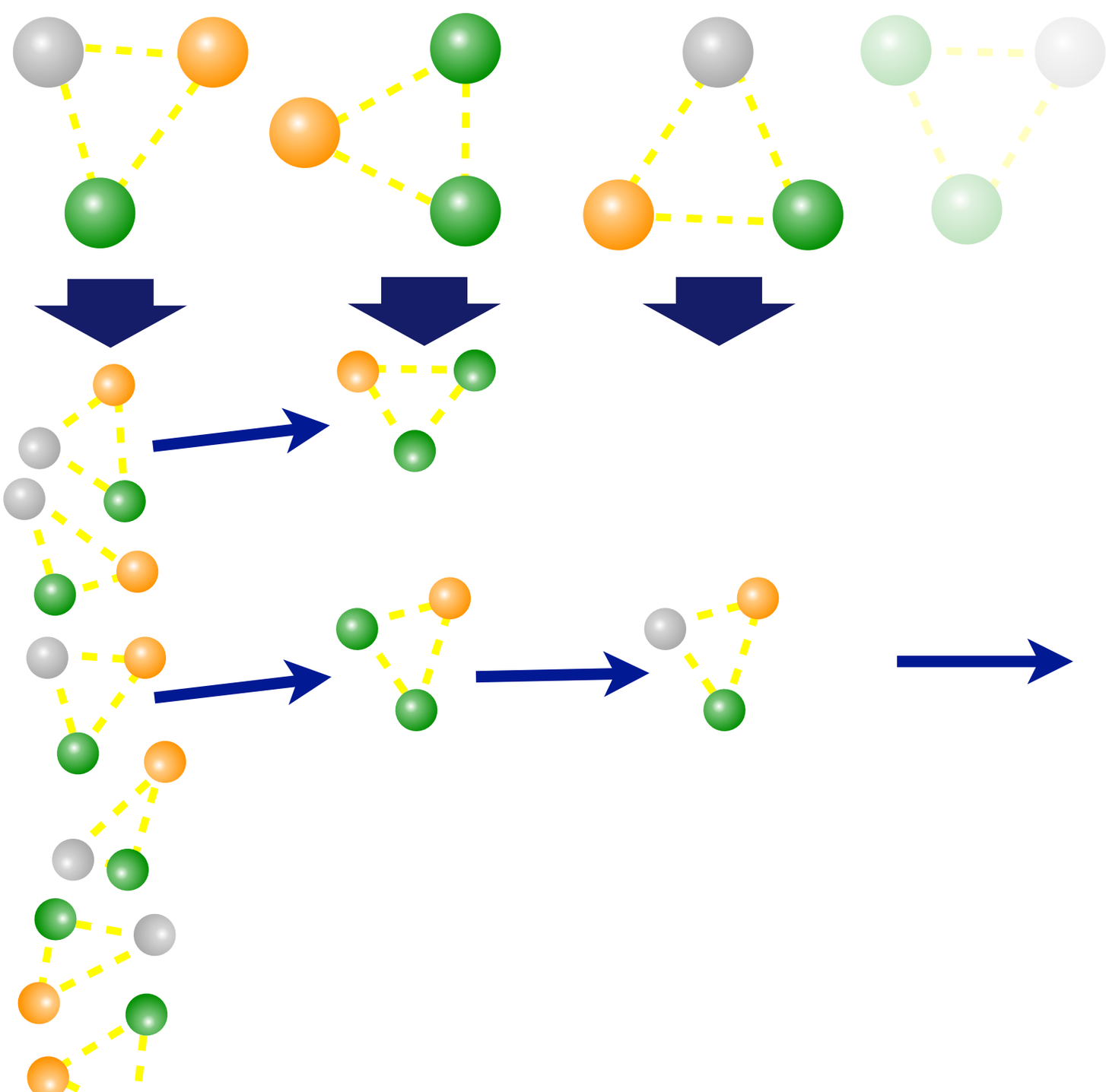
Efficient and Exact Pharmacophore Search

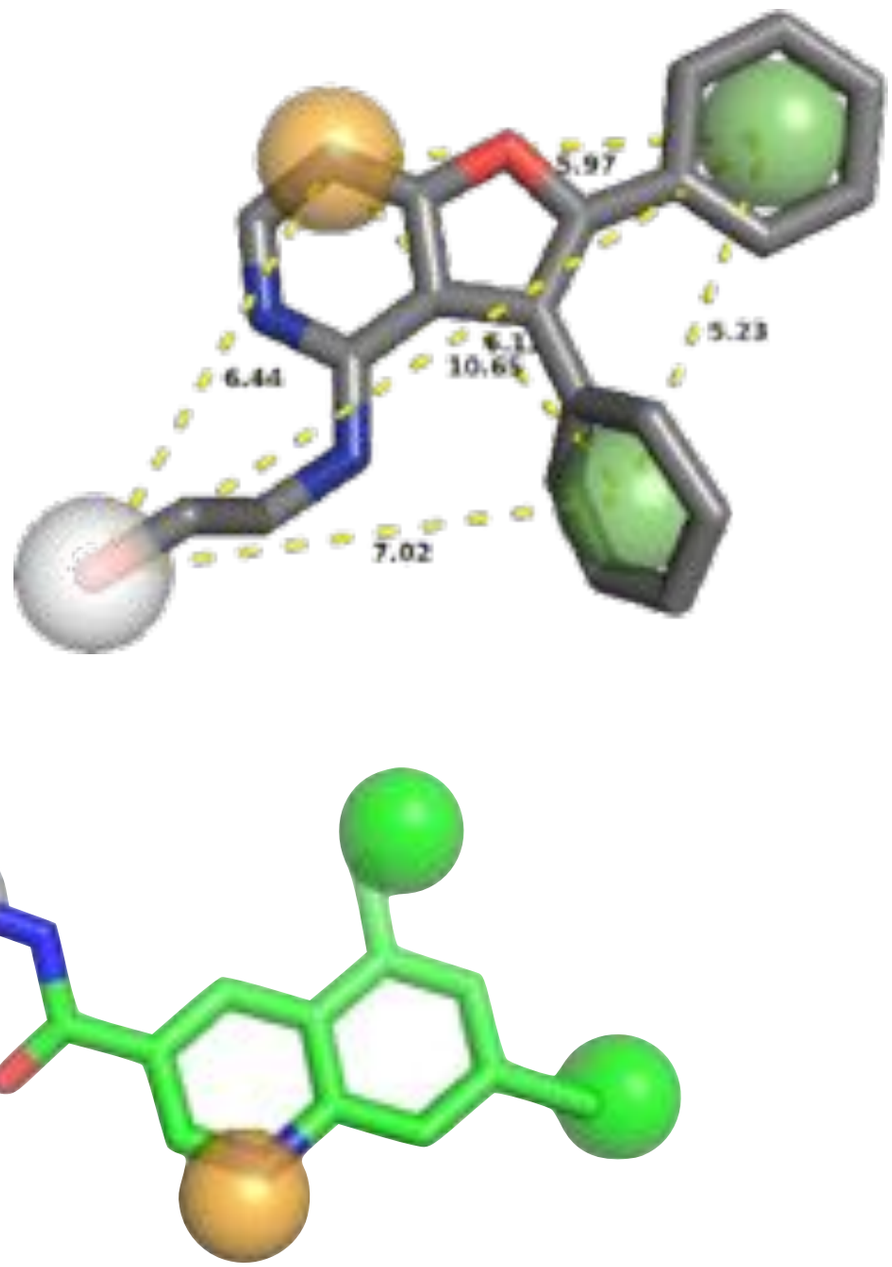




Pharmer

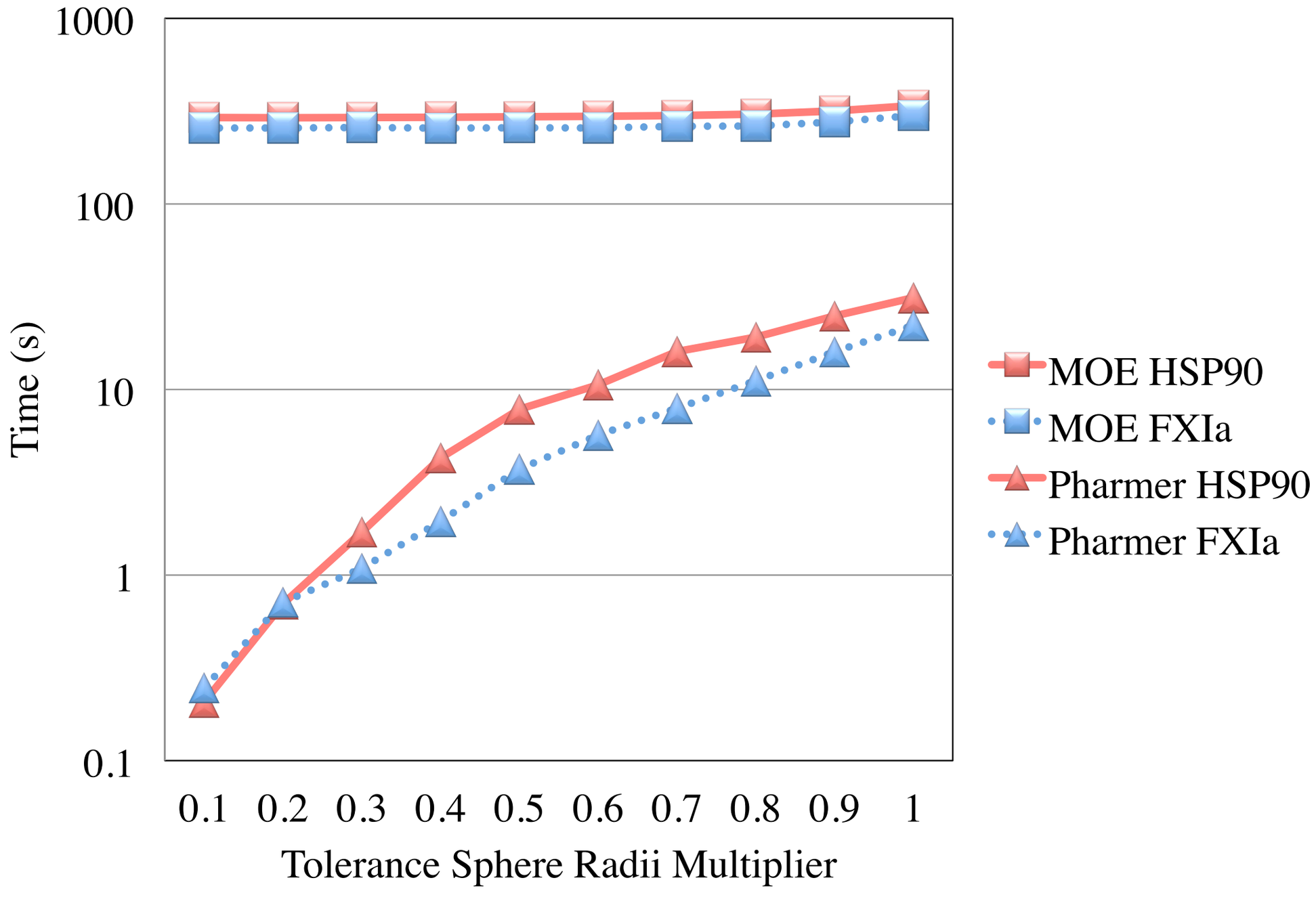
Efficient and Exact Pharmacophore Search

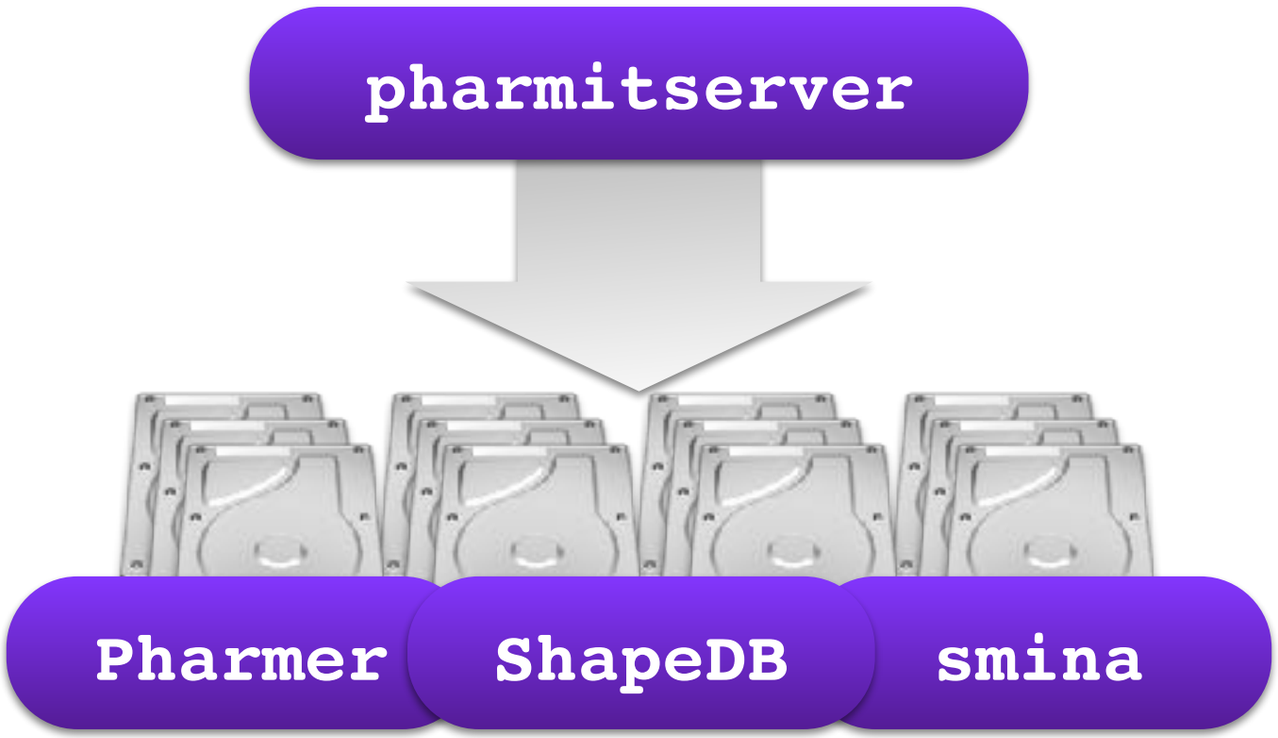




Pharmer

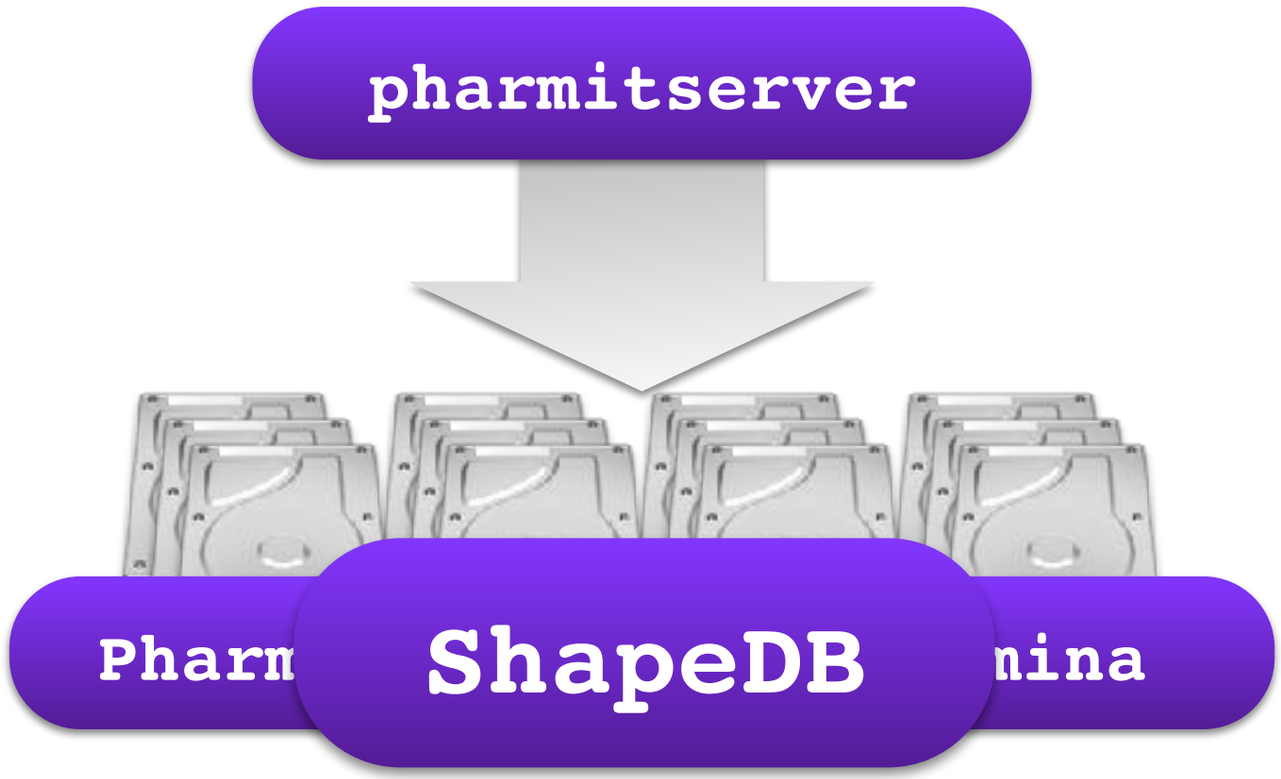
Efficient and Exact Pharmacophore Search





http://pharmit.sf.net

GPL/BSD License



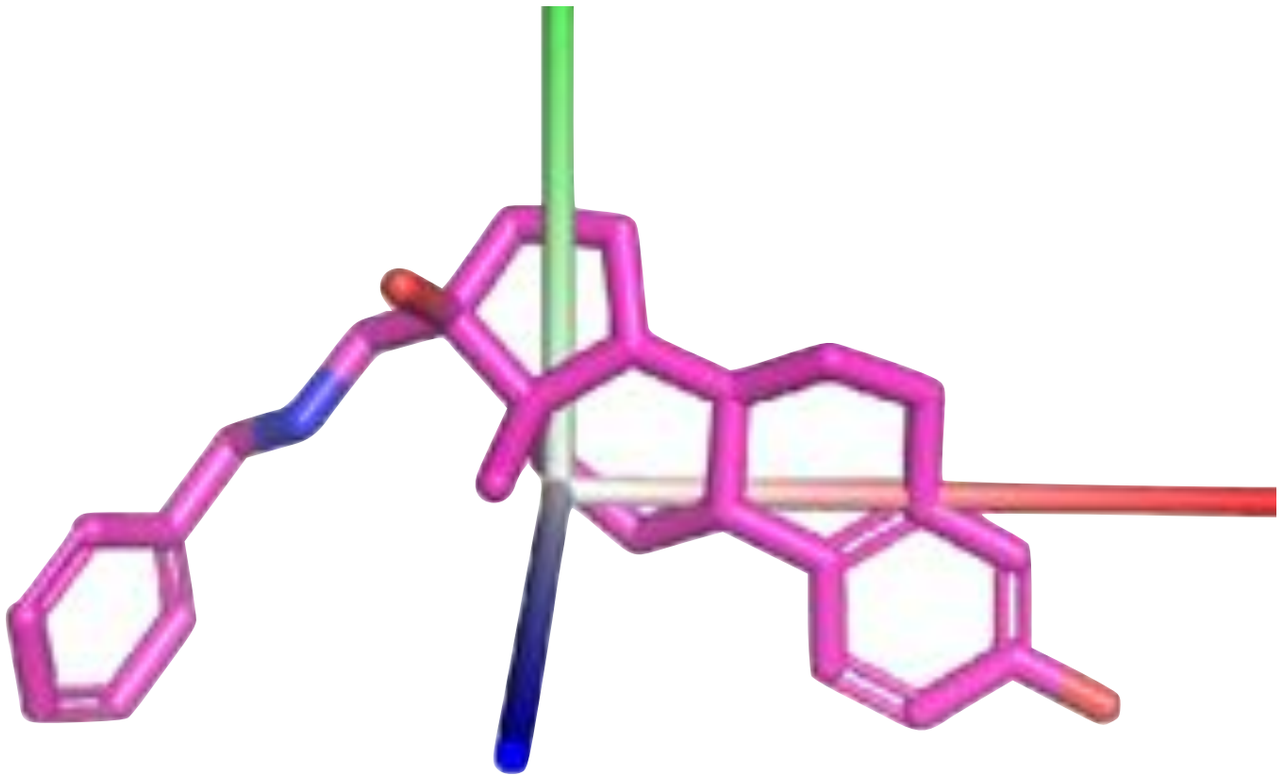
http://pharmit.sf.net

GPL/BSD License

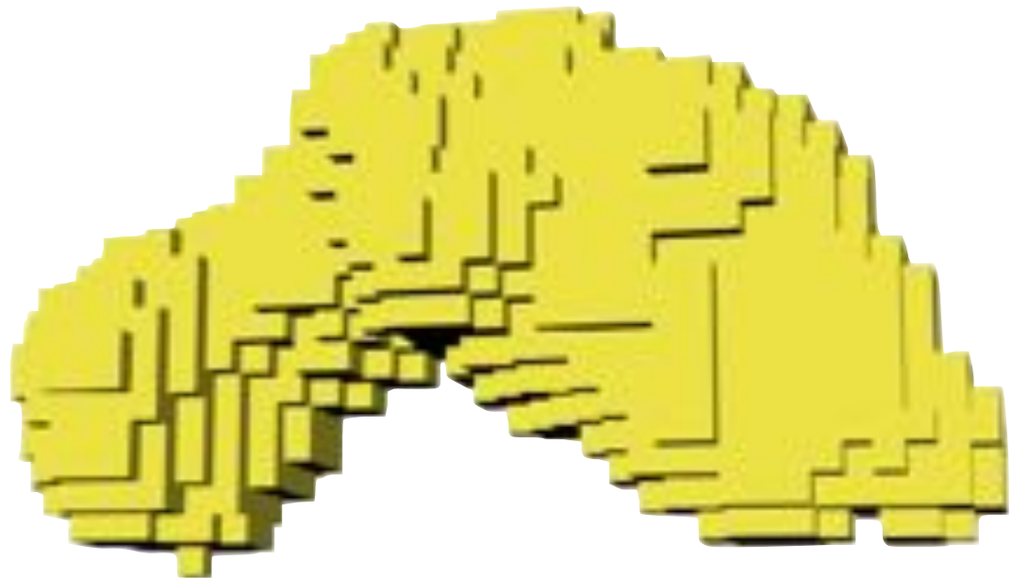
ShapeDB

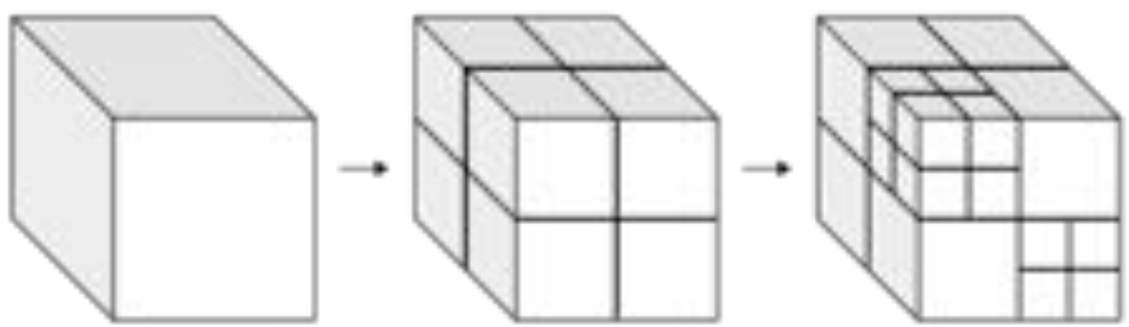
Indexed Search of Molecular Shapes

Align to Moments of Inertia



Voxelize





Oct-tree

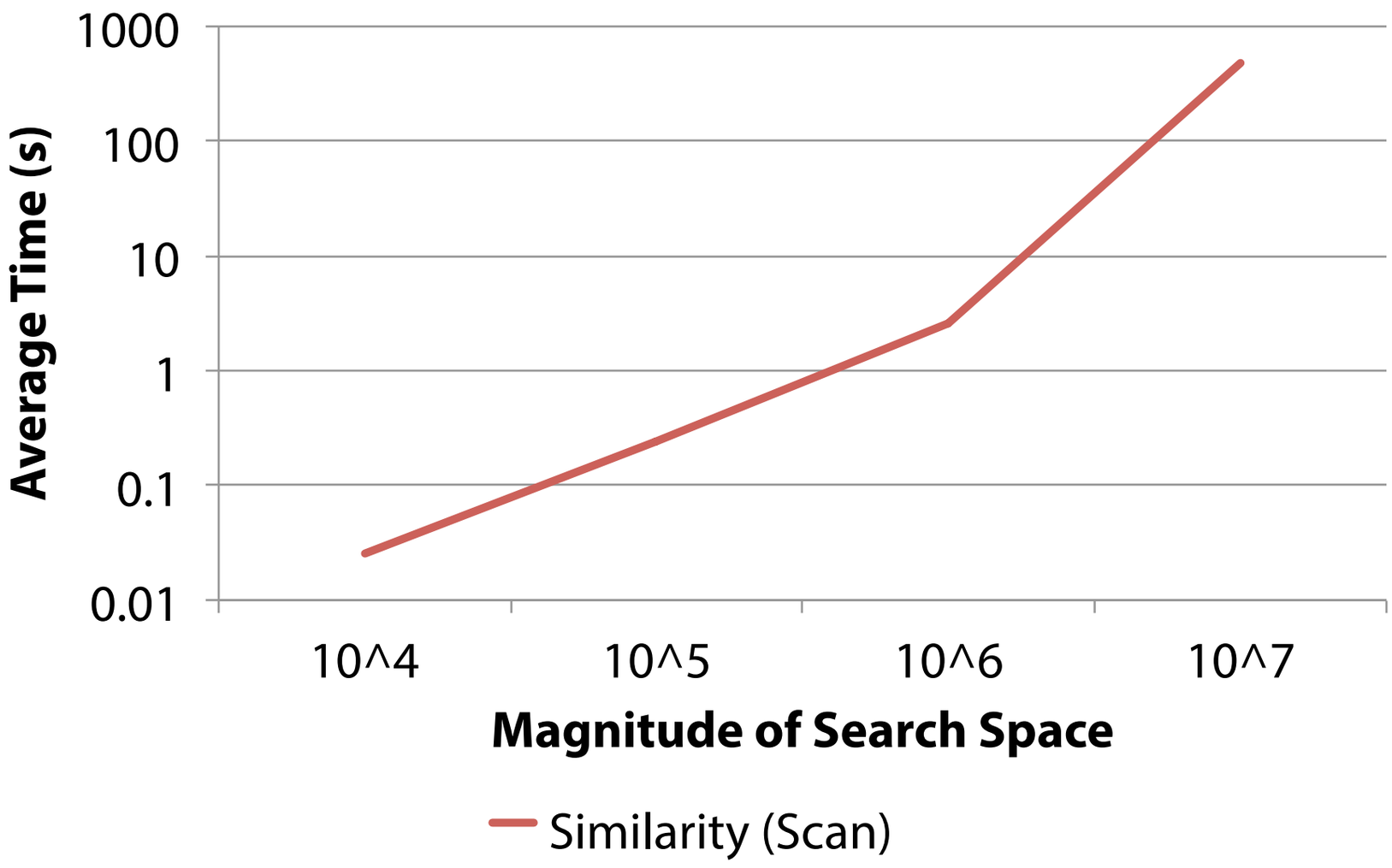
- Scales with Surface Area, not Volume
- Fast Intersection/Union Operations

ShapeDB

Indexed Search of Molecular Shapes

$$\delta(A, B) = \frac{A \cap B}{A \cup B}$$

Performance of Shape Similarity Search



ShapeDB

Indexed Search of Molecular Shapes

MIV

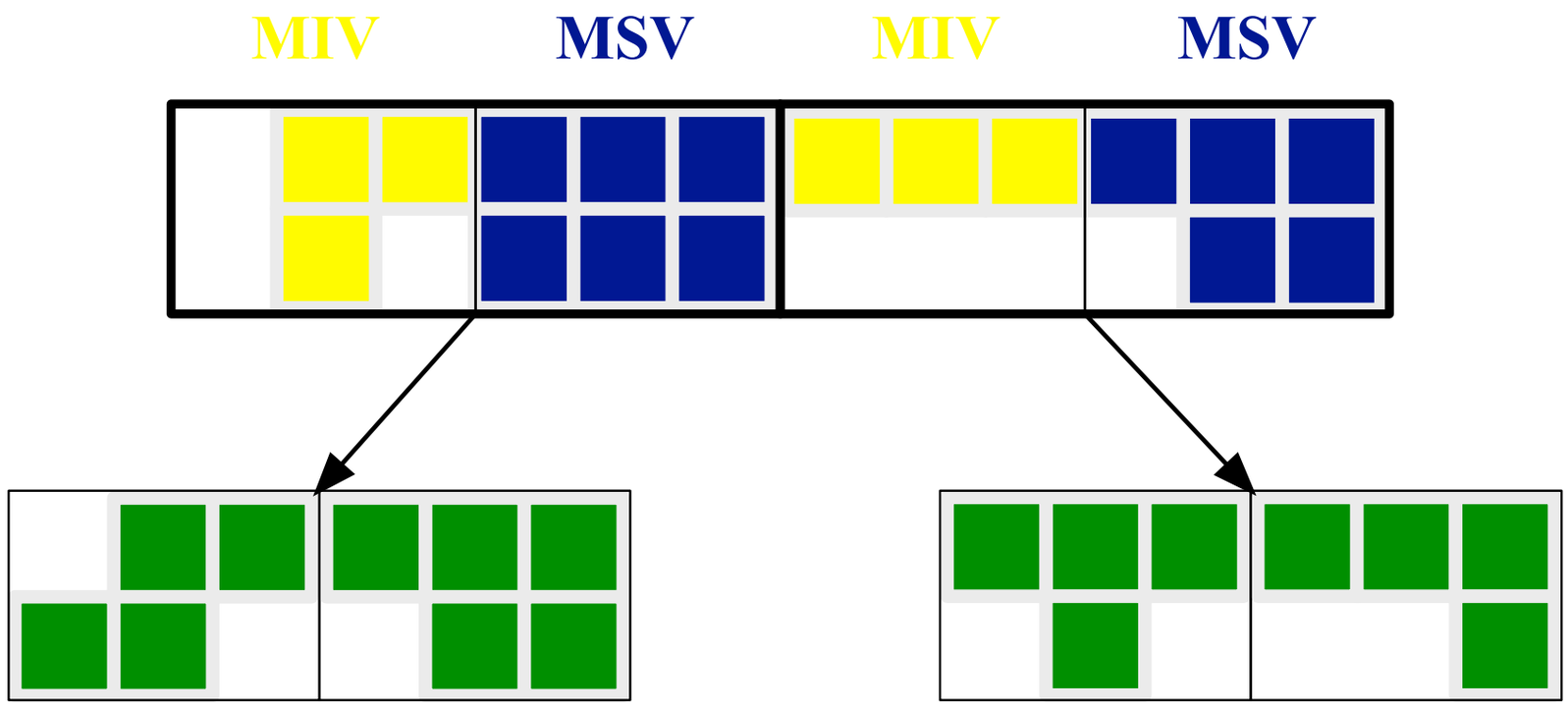
Maximum Included Volume

Intersection

MSV

Minimum Surrounding Volume

Union

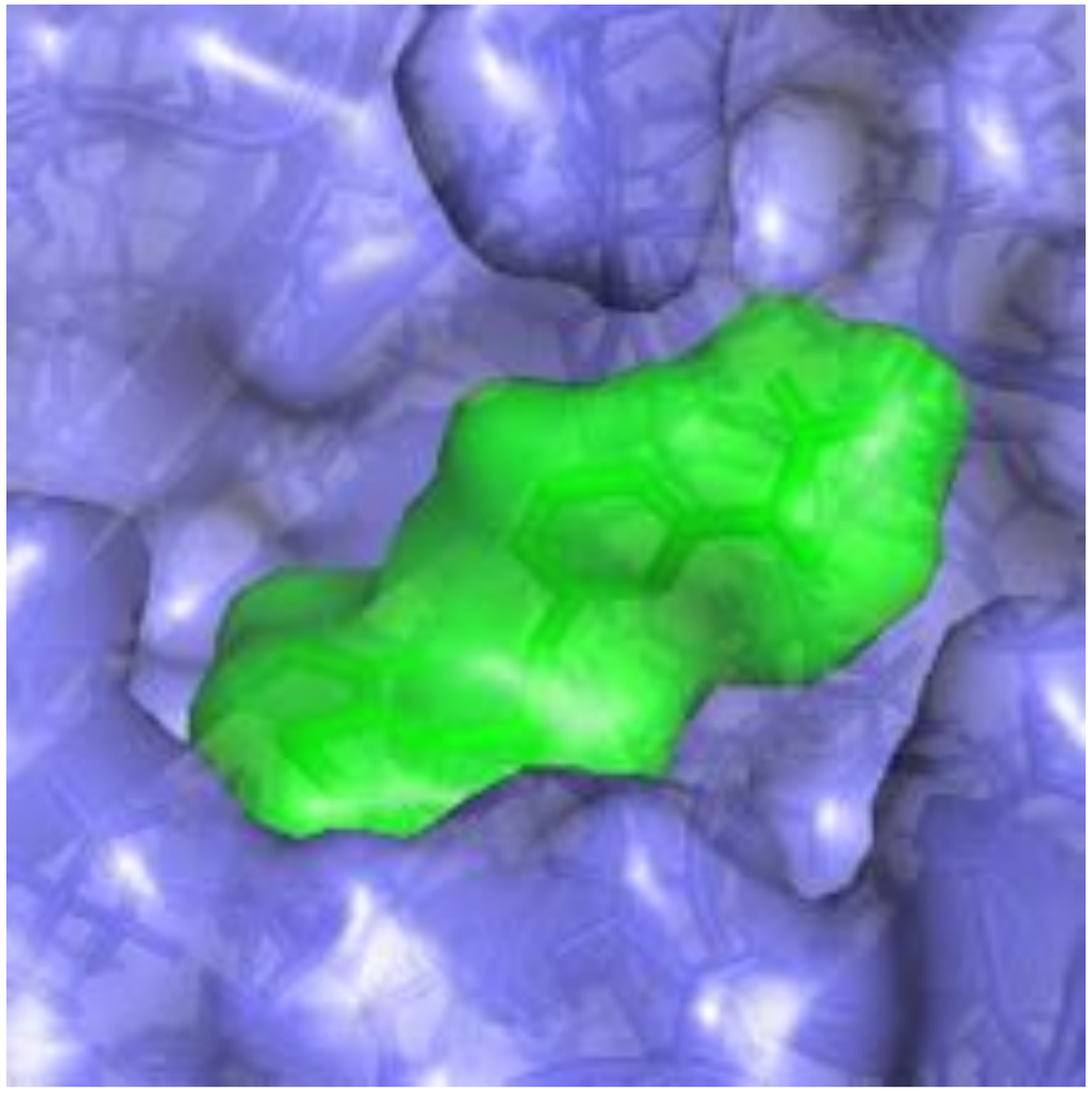


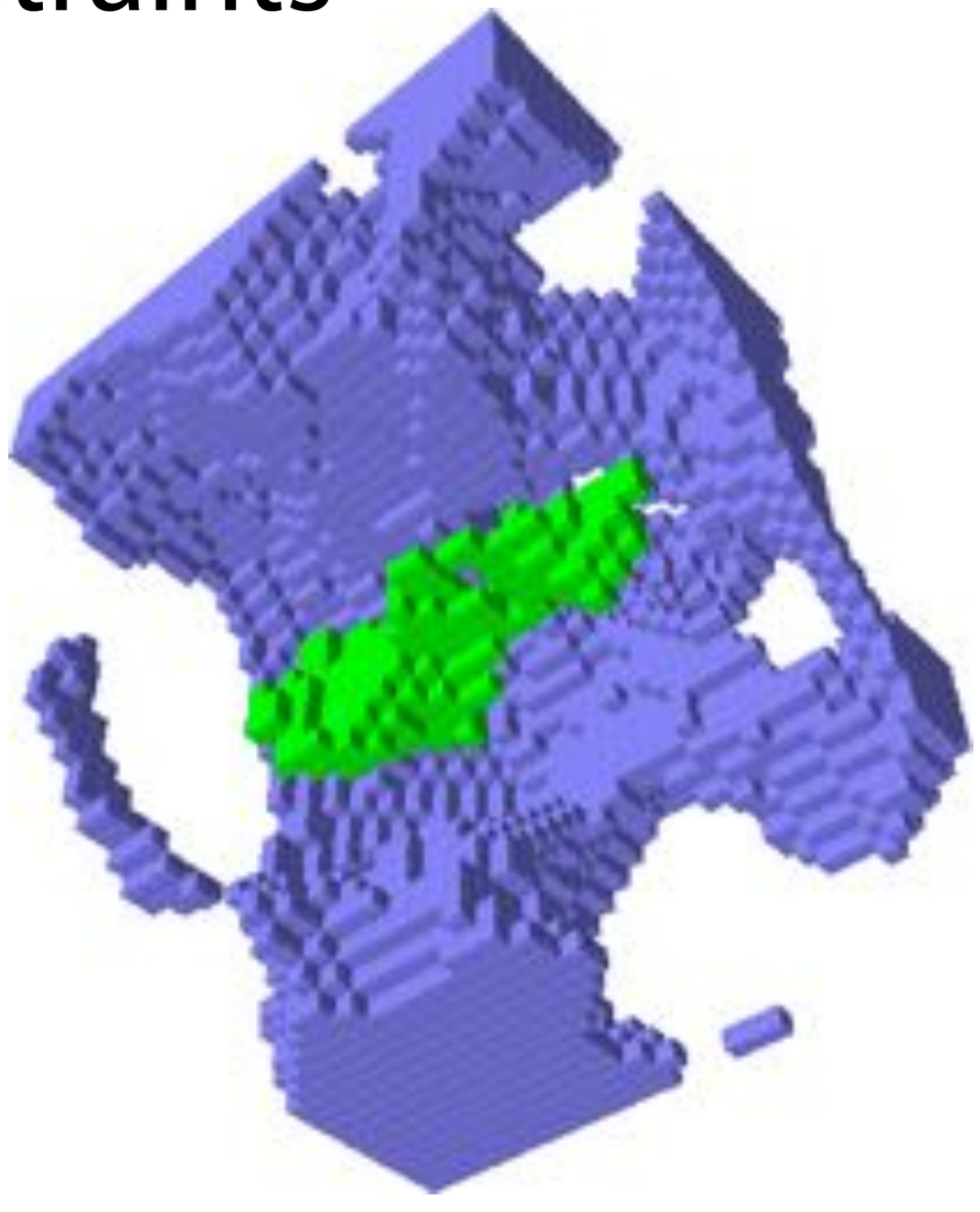
Matching and packing algorithm for efficient and effective initialization

Koes, D. R., & Camacho, C. J. (2014). Shape-based virtual screening with volumetric aligned molecular shapes. *J Comput Chem, 35*(25), 1824-1834. doi:10.1002/jcc.23690

Koes, D., & Camacho, C. (2014). Indexing volumetric shapes with matching and packing. *Knowledge and Information Systems*, 1-24. doi:10.1007/s10115-014-0729-z

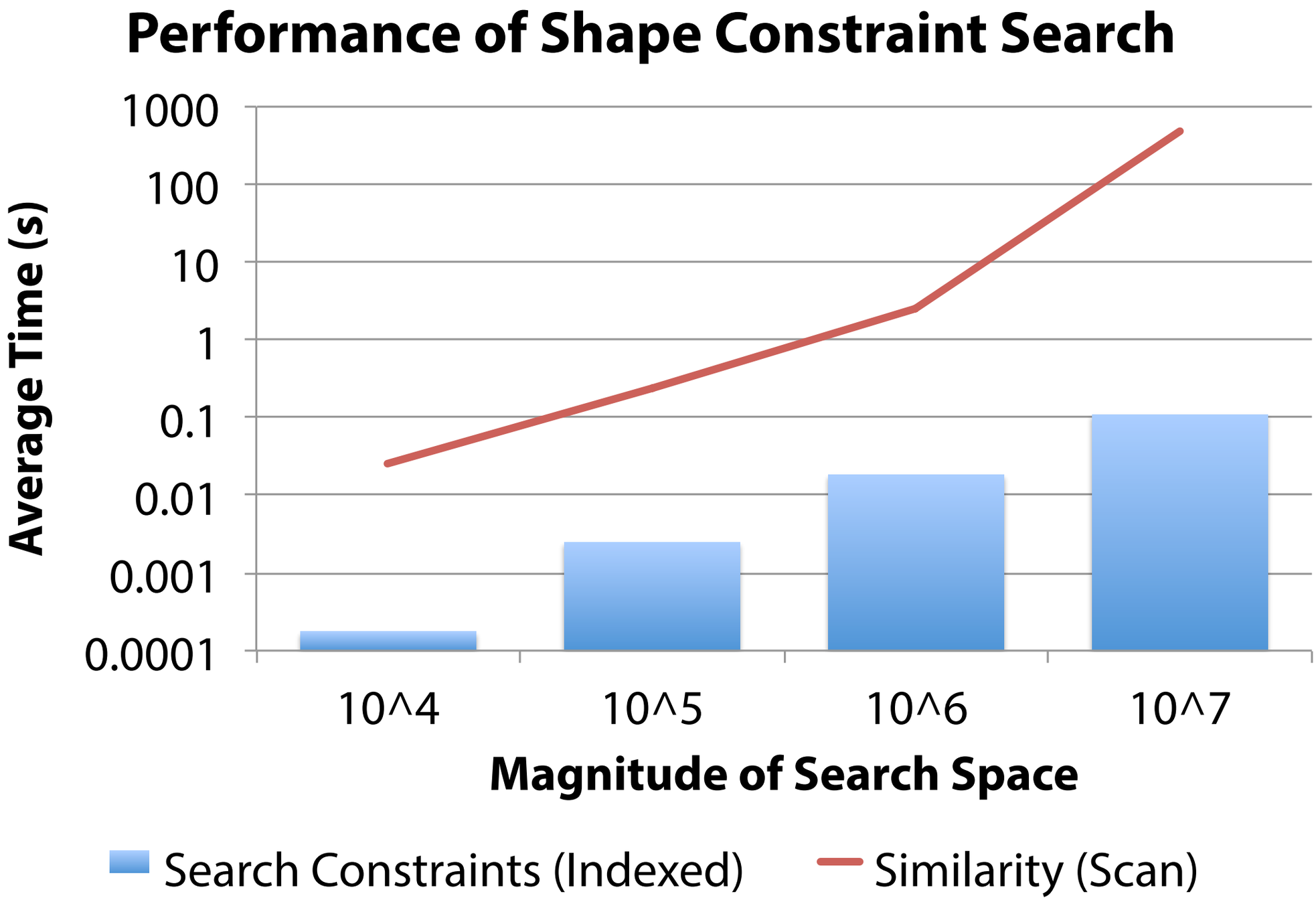
Shape Constraints

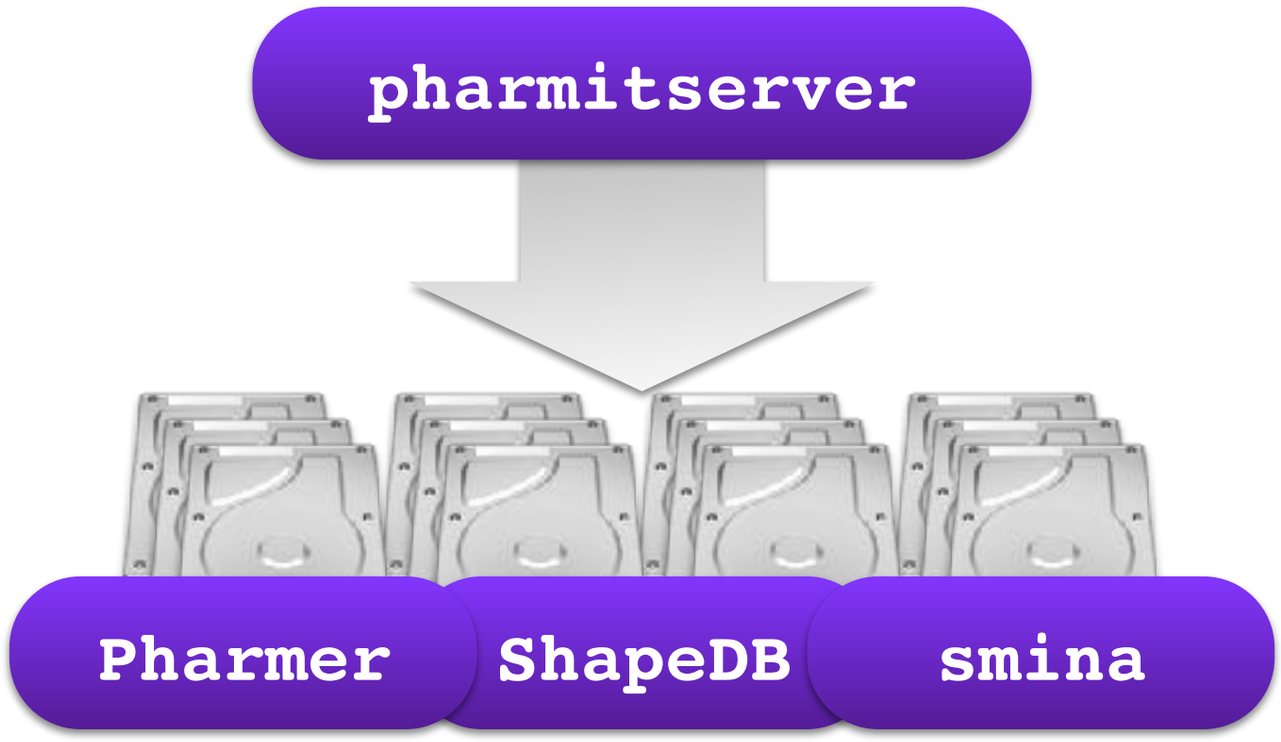




ShapeDB

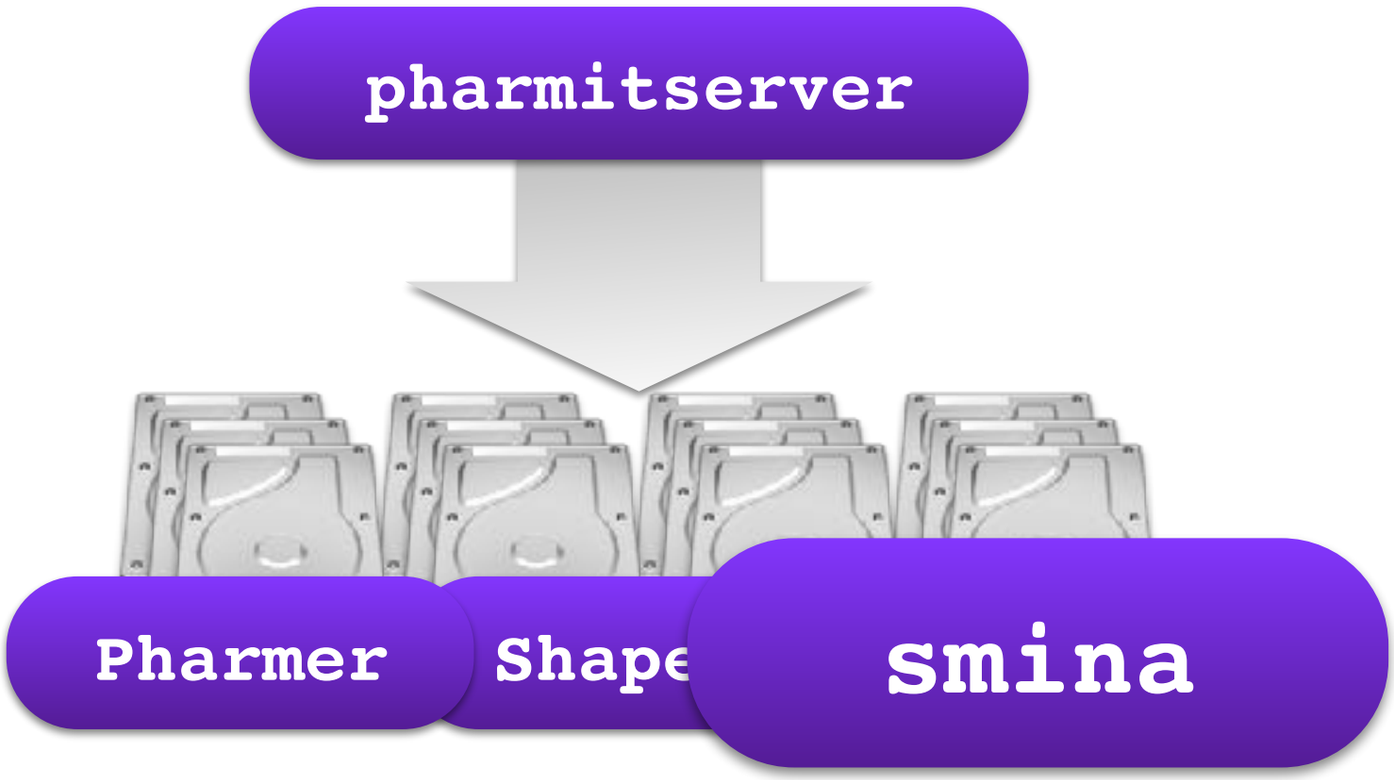
Indexed Search of Molecular Shapes





http://pharmit.sf.net

GPL/BSD License

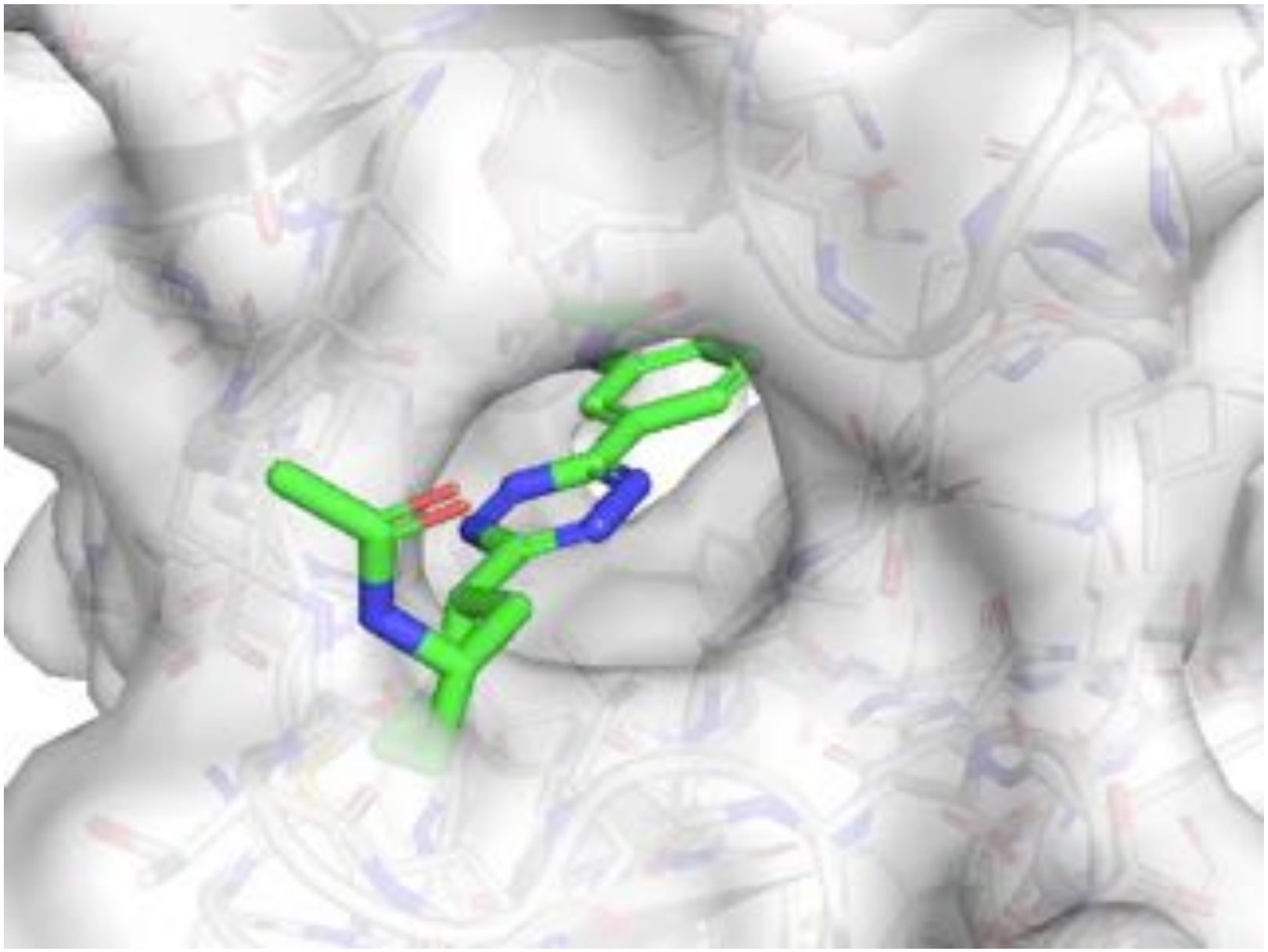


http://pharmit.sf.net

GPL/BSD License

smina

Scoring and Minimization with AutoDock Vina

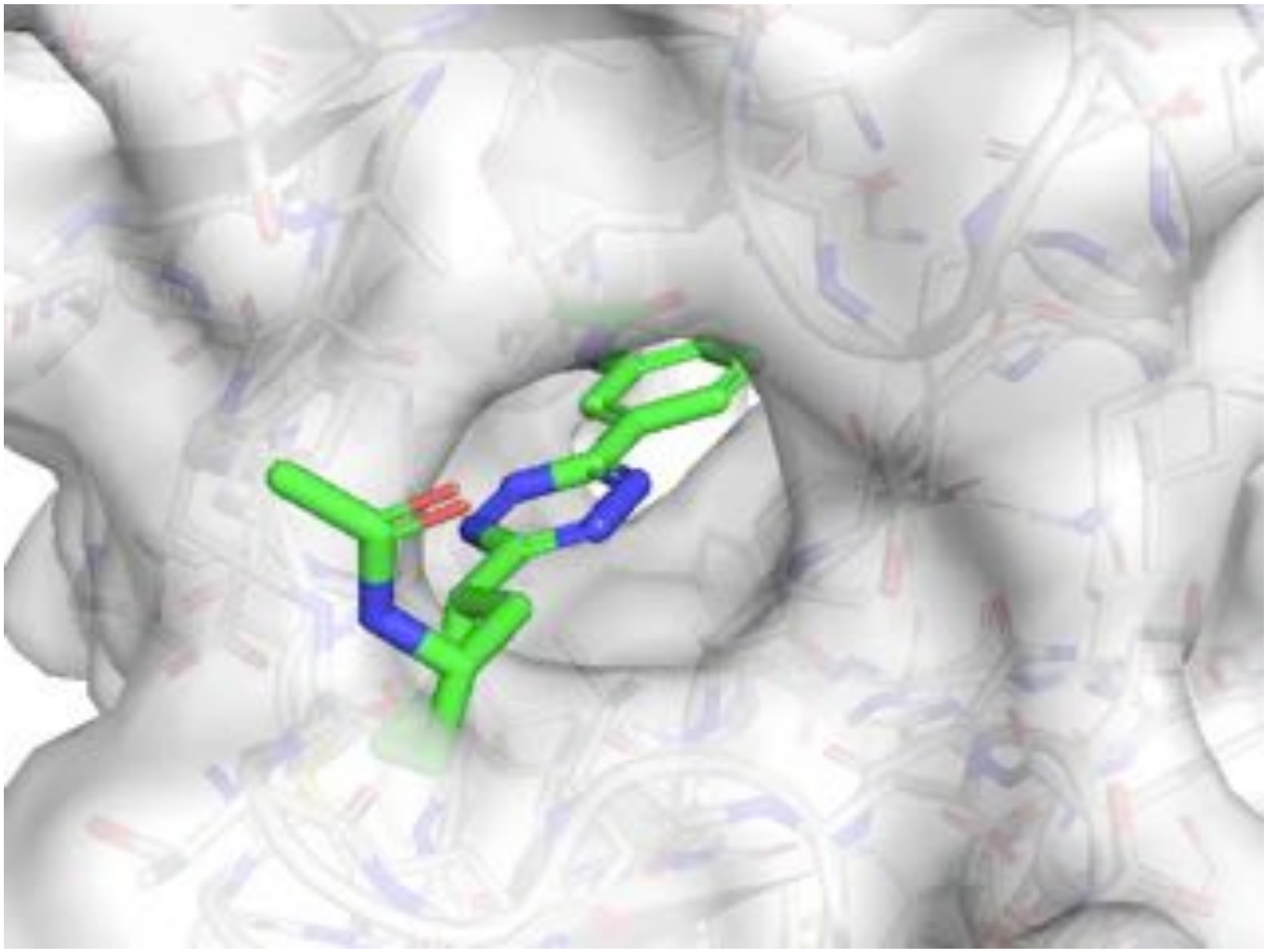


- Improved energy minimization
- Facilities for scoring function development
- Custom format with precomputed torsion tree
- Implementation of server protocol

Koes, D. R., Baumgartner, M. P., & Camacho, C. J. (2013). Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. *Journal of Chemical Information and Modeling, 53*(8), 1893-1904. doi:10.1021/ci300604z

smina

Scoring and Minimization with AutoDock Vina

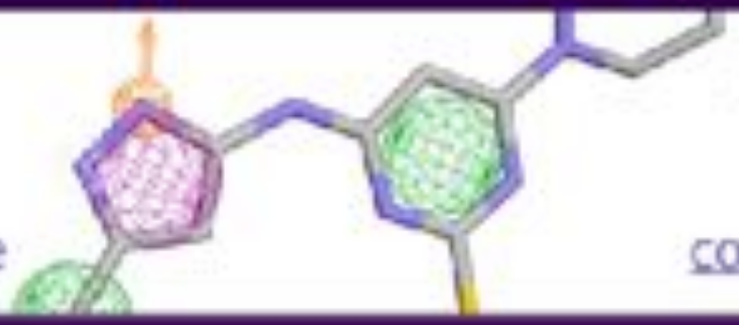


- Improved energy minimization
- Facilities for scoring function development
- Custom format with precomputed torsion tree
- Implementation of server protocol

Koes, D. R., Baumgartner, M. P., & Camacho, C. J. (2013). Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. *Journal of Chemical Information and Modeling, 53*(8), 1893-1904. doi:10.1021/ci300604z

pharmit beta

interactive exploration of chemical space



code

help

contact

search

virtual screening in your browser

enter pharmit search

start from PCB:

binding site waters:

submit

examples

create

submit your own chemical libraries

log in to manage libraries

email:

password:

log in

register new account

log in as guest

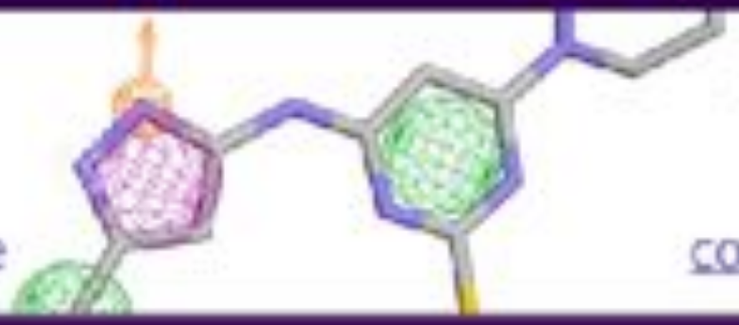
pharmit currently has

5 built-in libraries containing 1,039,705,461 conformations of 75,076,775 compounds and

7 publicly accessible user-contributed libraries containing

pharmit beta

interactive exploration of chemical space



code

help

contact

search

virtual screening in your browser

enter pharmit search

start from PCB:

binding site waters:

submit

examples

create

submit your own chemical libraries

log in to manage libraries

email:

password:

log in

register new account

log in as guest

pharmit currently has

5 built-in libraries containing 1,039,705,461 conformations of 75,076,775 compounds and

7 publicly accessible user-contributed libraries containing

Acknowledgements

Group Members

Jocelyn Sunseri
Matt Ragoza
Josh Hochuli
Roosha Mandal
Alec Helbling
Lily Turner
Aaron Zheng
Sara Amato
Lily Turner
Aaron Zheng
Gibran Biswas

Department of Computational and Systems Biology

NIGMS R01GM108340

http://pharmit.csb.pitt.edu

@david_koes

http://bits.csb.pitt.edu