
Inseguro | http://www.way2drug.com/dr/symp_bcadd_2022.php
Way2Drug
PREDICTIVE SERVICES
Understanding Chemical-Biological Interactions
XXVIII Symposium on Bioinformatics and Computer-Aided Drug Discovery
SYMPOSIUM
COMMITTEE
PROGRAM
PROCEEDINGS
AWARDS
REGISTRATION/LOGIN
IYBSSD2022
International Year of Basic Sciences for Sustainable Development
Moscow, Russia
24-26 May, 2022 | VIRTUAL
IBMC
Institute of Biomedical Chemistry
XXVIII Symposium on Bioinformatics and Computer-Aided Drug Discovery
Emerging Challenges and Opportunities for In Silico Drug Discovery

Keynote Lecture
Harnessing Machine Learning for Drug Development

Dr. Walter F. de Azevedo, Jr.
azevedolab.net

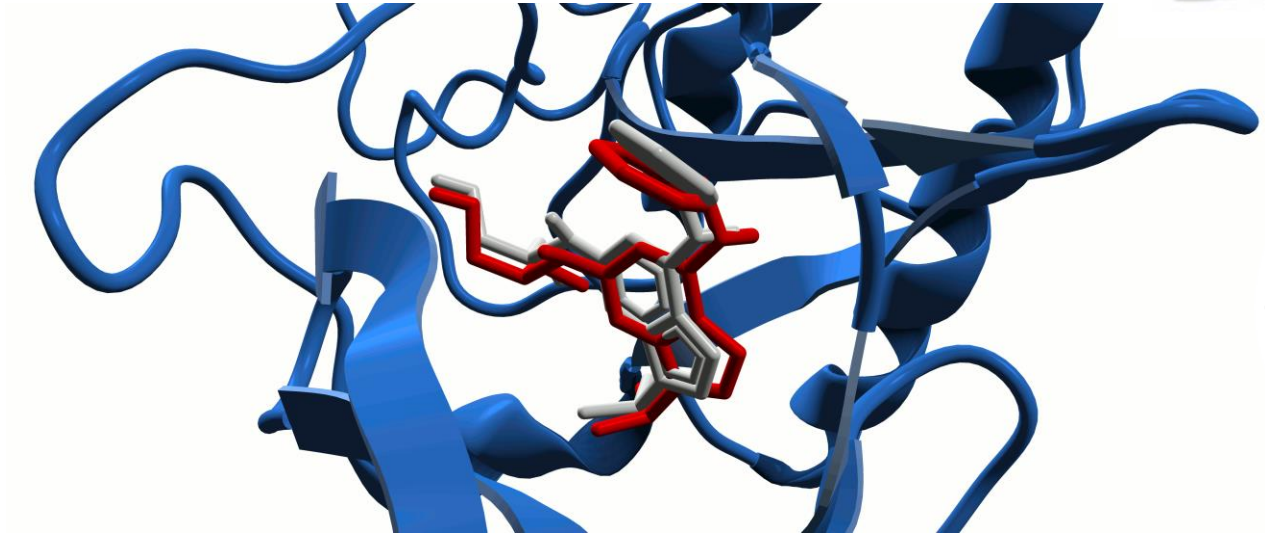
walter@azevedolab.net

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
Porto Alegre/RS
Brazil

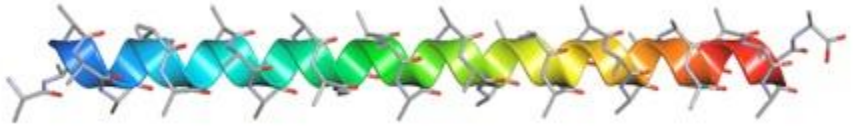
- Overview
- Chemical Space
- Protein Space
- Scoring Function Space
- Cyclin-Dependent Kinase 2
- Taba
- SAnDReS

Overview

AutoDock 4



ArgusLab



Overview

$$V = W_{vdw} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) +$$

$$W_{HBond} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right)$$

$$W_{elec} \sum_{i,j} \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} +$$

$$W_{sol} \sum_{i,j} (S_i v_j + S_j v_i) e^{\left(-r_{ij}^2 / 2\sigma^2\right)} + W_{tor} N_{Tor}$$

AutoDock 4

Source: Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem. 2009; 30(16):2785-91.

Overview

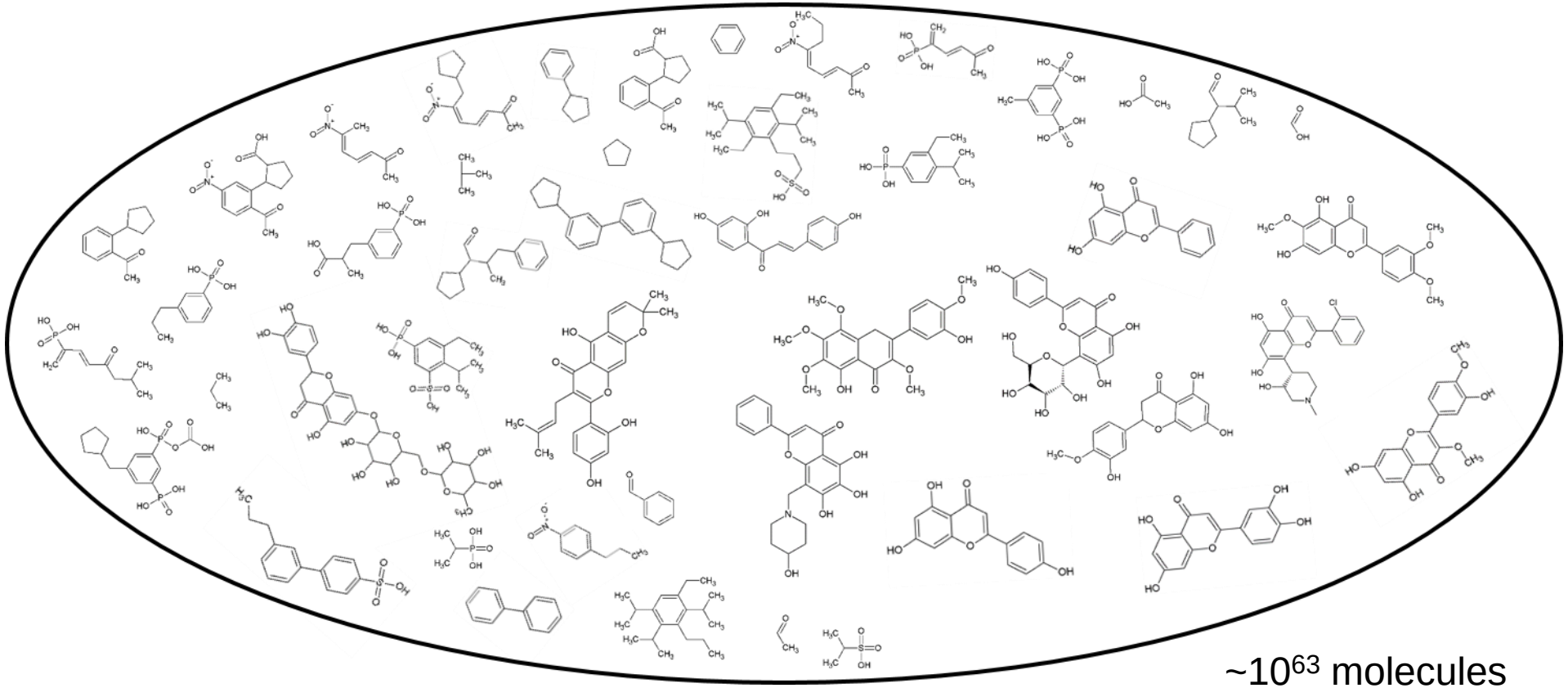
Regression model prediction

$$\hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \ldots + \theta_n x_n$$

$\hat{y}$ is the predicted value
θ_j is the j^{th} model parameter
x_i is the i^{th} feature value
n is the number of features

Source: Gérpon A. Hands On Machine Learning with Scikit-Learn, Keras & TensorFlow. 2nd Ed. Sebastopol: O'Reilly Media, Inc., 2019. 819 p.

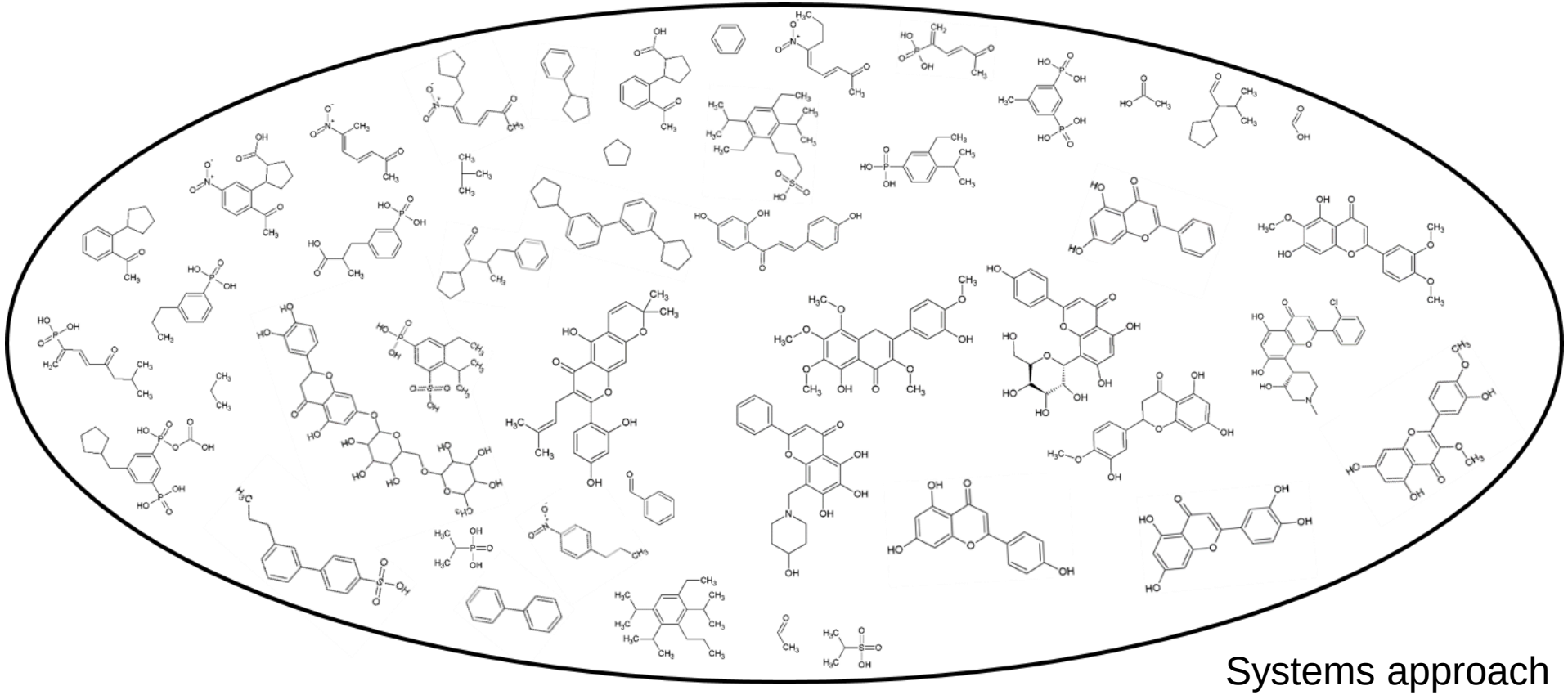
Chemical Space



~10^{63} molecules

Source: Bohacek RS, McMartin C, Guida WC. The art and practice of structure-based drug design: a molecular modeling perspective. Med Res Rev. 1996; 16(1):3–50.

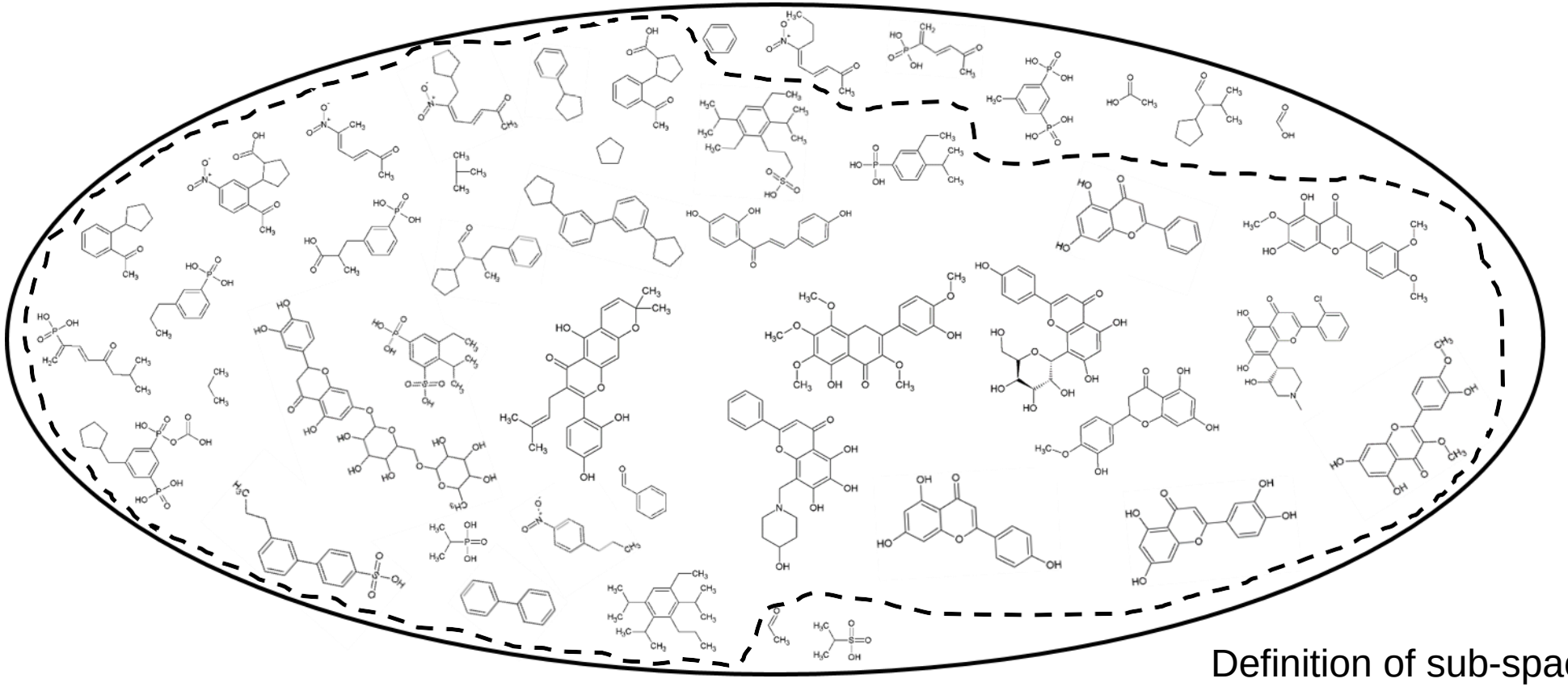
Chemical Space - Advantages



Systems approach

Source: Bohacek RS, McMartin C, Guida WC. The art and practice of structure-based drug design: a molecular modeling perspective. Med Res Rev. 1996; 16(1):3–50.

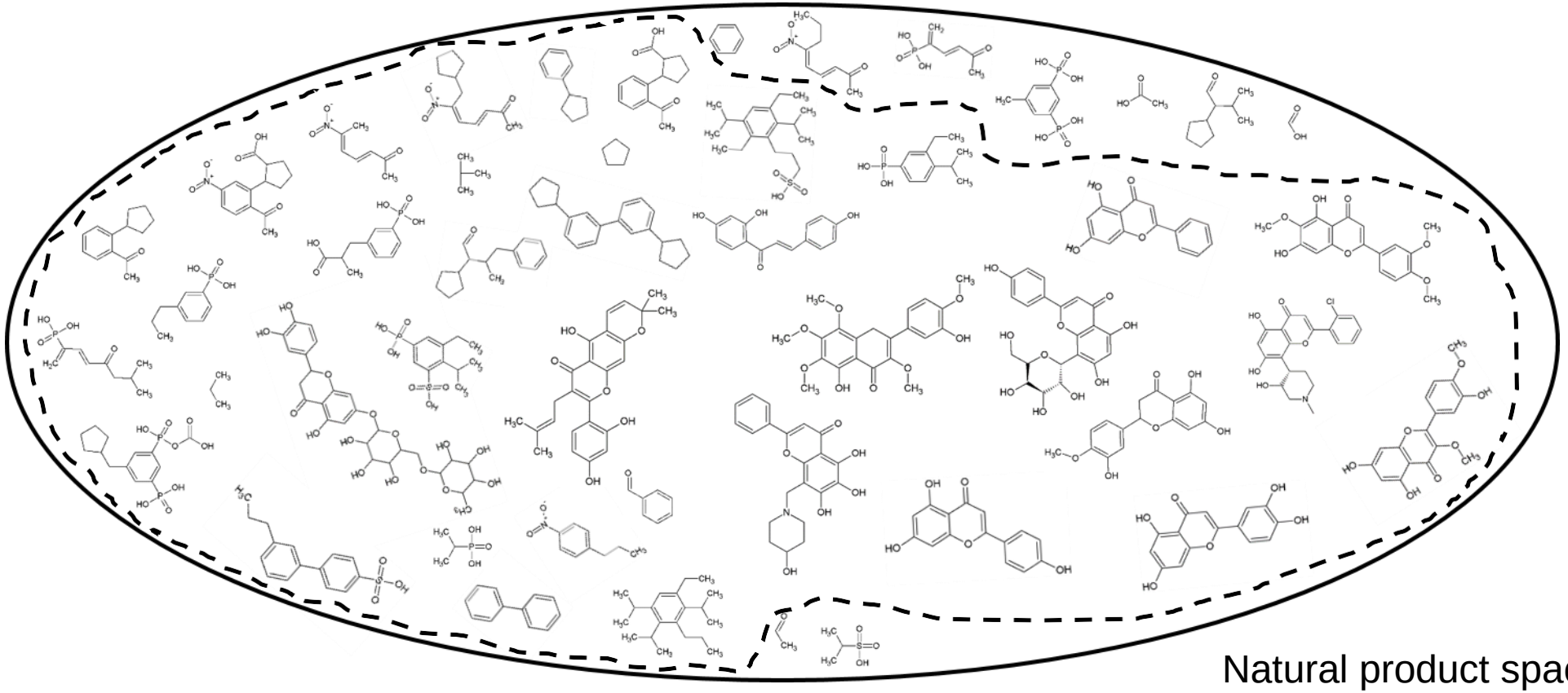
Chemical Space - Advantages



Definition of sub-spaces

Source: Bohacek RS, McMartin C, Guida WC. The art and practice of structure-based drug design: a molecular modeling perspective. Med Res Rev. 1996; 16(1):3–50.

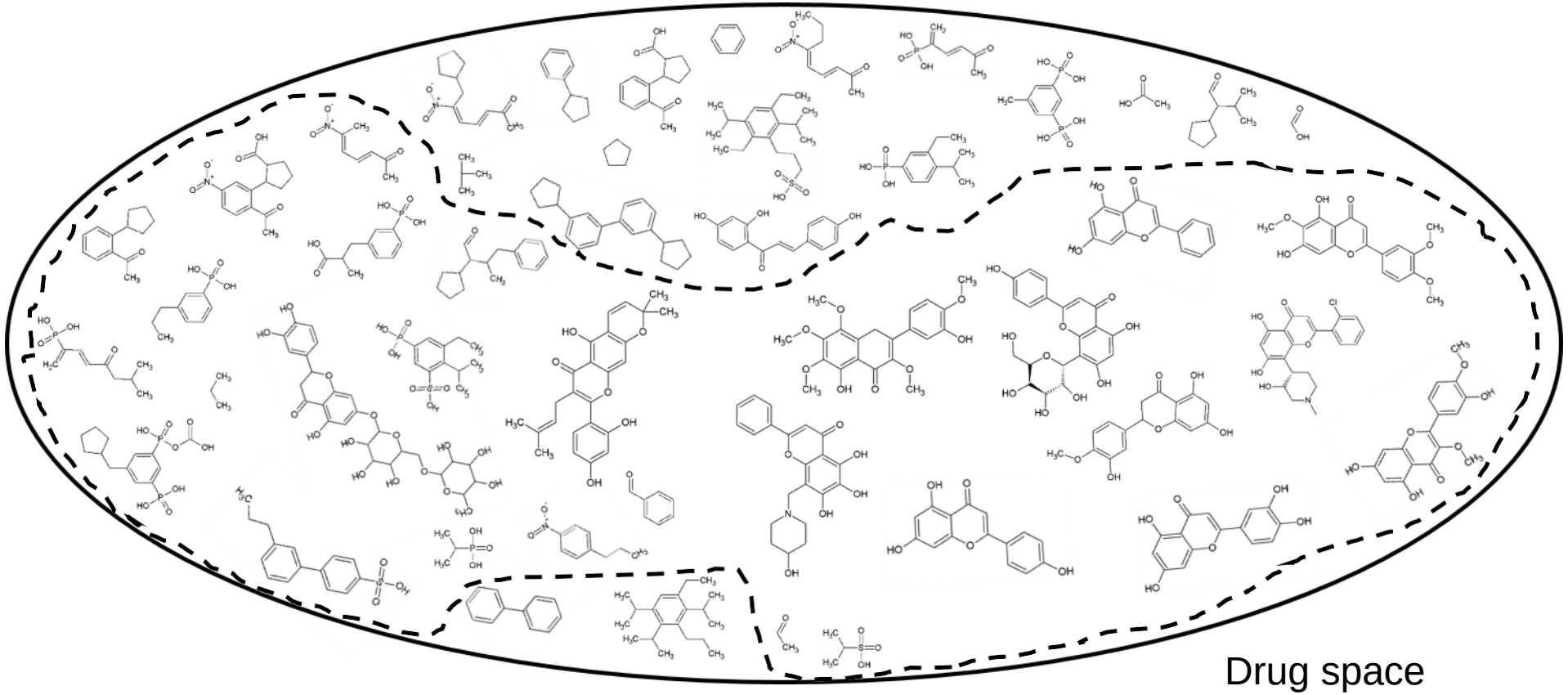
Chemical Space - Advantages



Natural product space

Source: Bohacek RS, McMartin C, Guida WC. The art and practice of structure-based drug design: a molecular modeling perspective. Med Res Rev. 1996; 16(1):3–50.

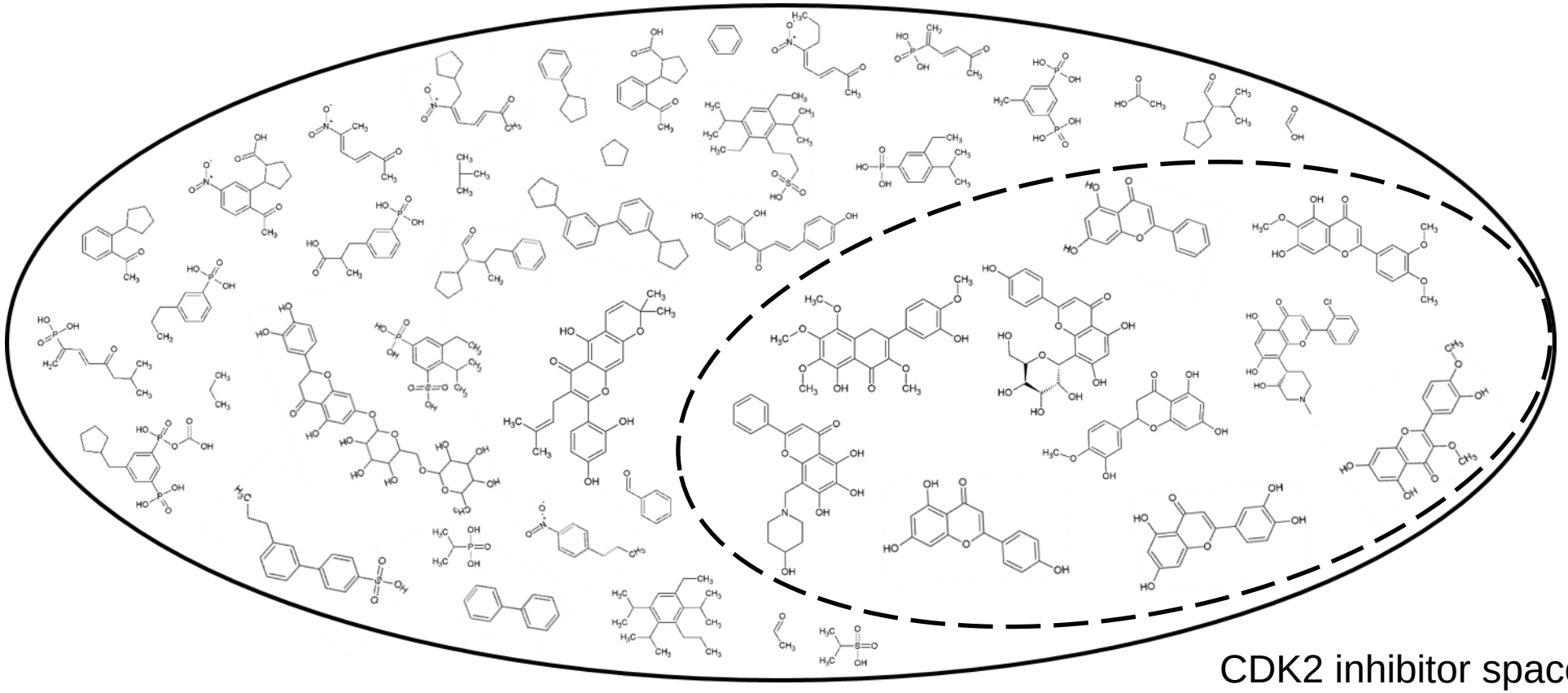
Chemical Space - Advantages



Drug space

Source: Bohacek RS, McMartin C, Guida WC. The art and practice of structure-based drug design: a molecular modeling perspective. Med Res Rev. 1996; 16(1):3–50.

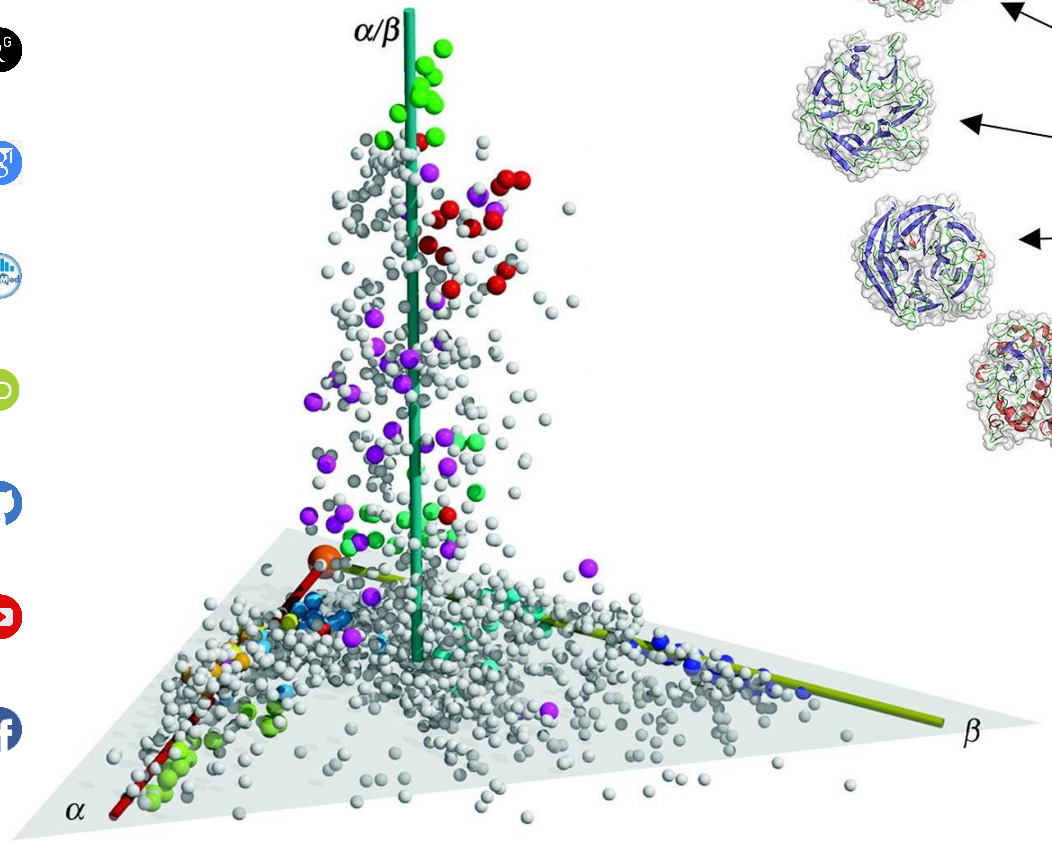
Chemical Space - Advantages

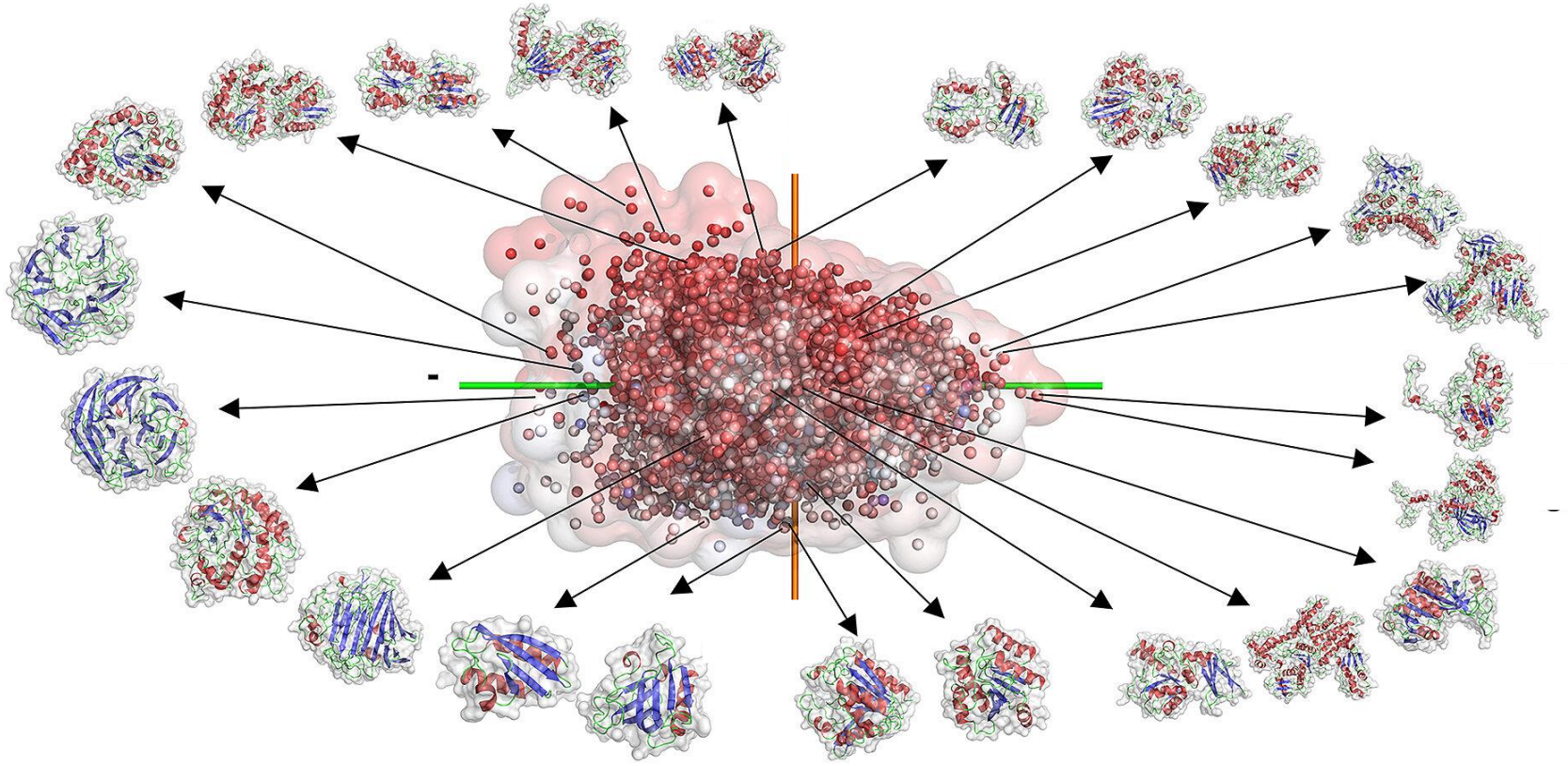


CDK2 inhibitor space

Source: Bohacek RS, McMartin C, Guida WC. The art and practice of structure-based drug design: a molecular modeling perspective. Med Res Rev. 1996; 16(1):3–50.

Protein Space

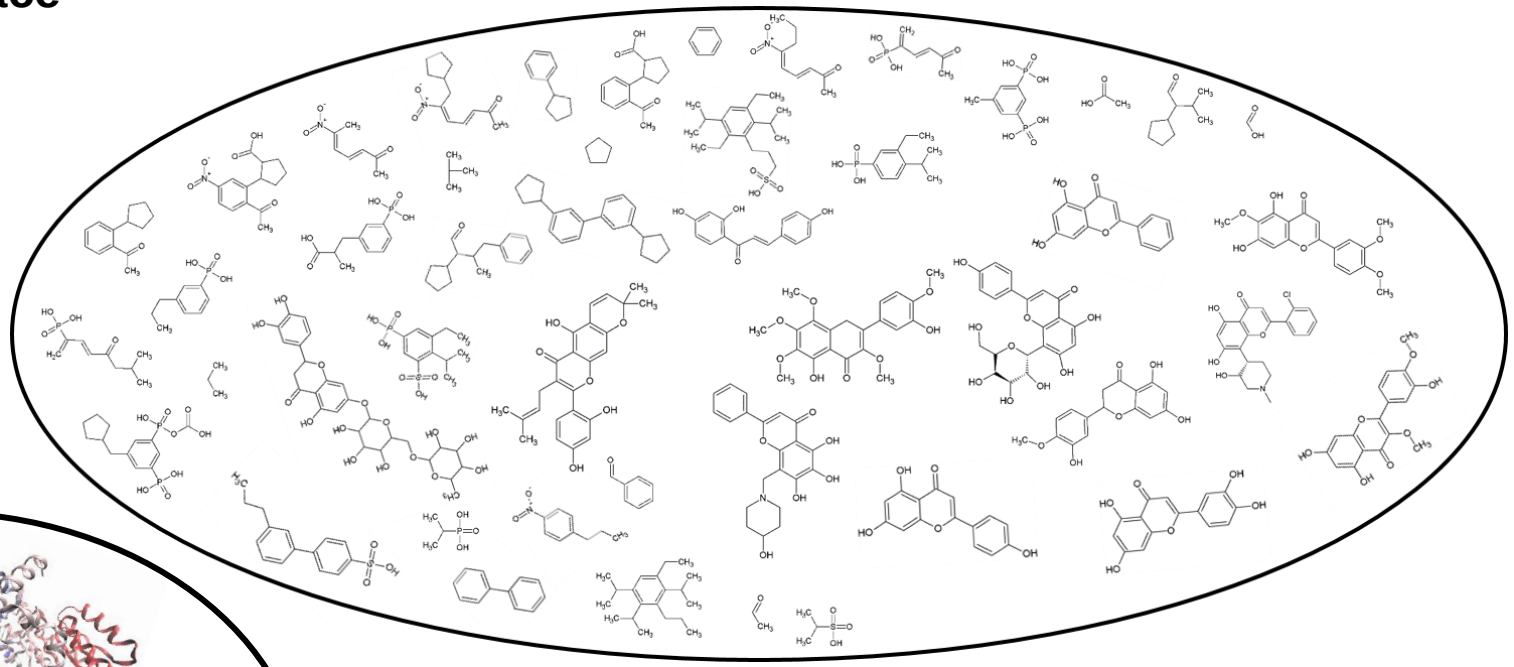


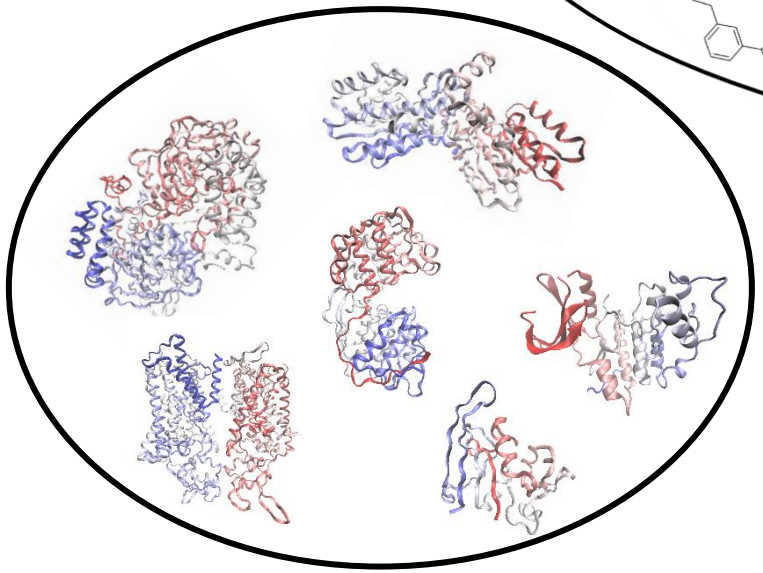


Source: Han X, Sit A, Christoffer C, Chen S, Kihara D. A global map of the protein shape universe. PLoS Comput Biol. 2019; 15(4):e1006969.

Scoring Function Space

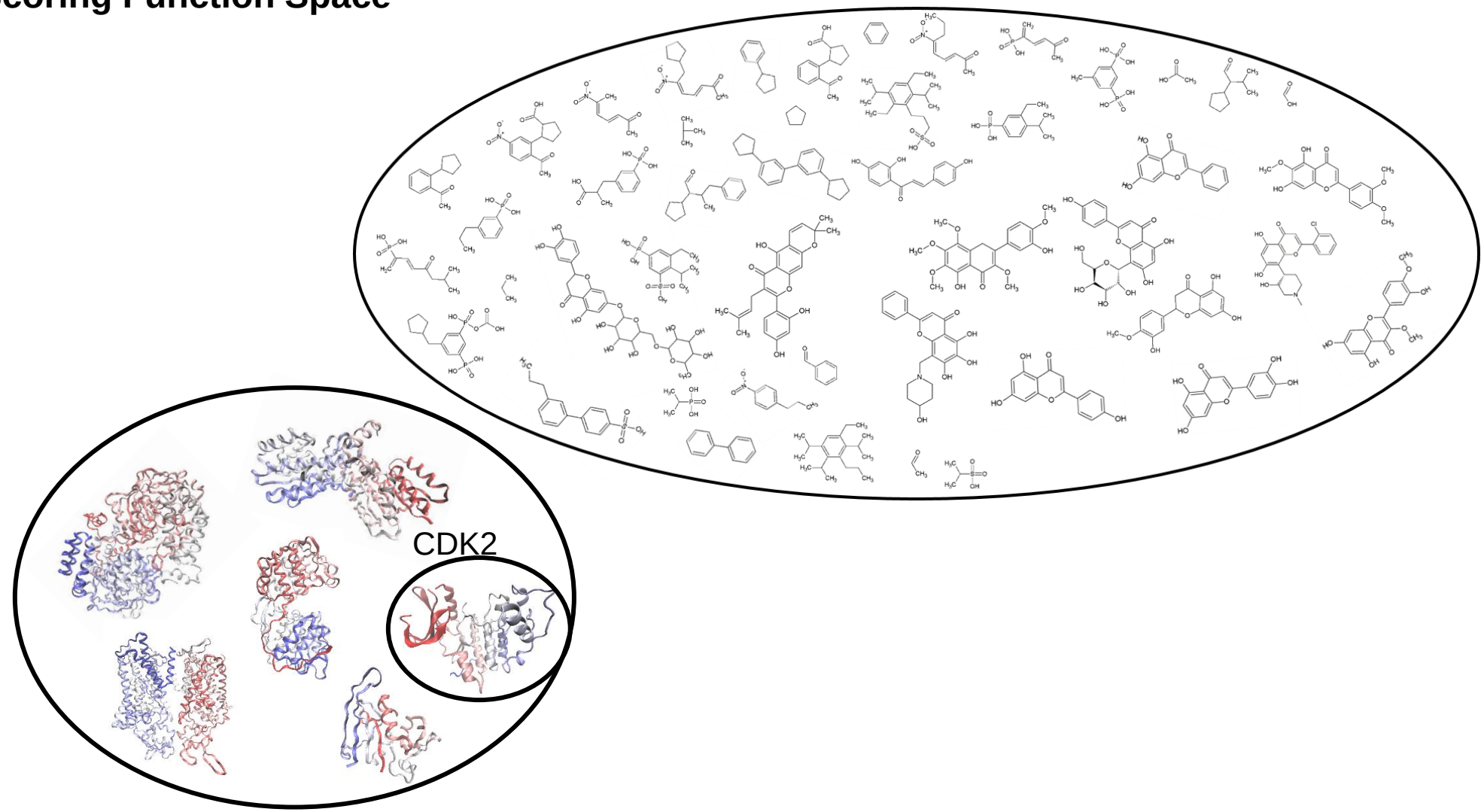
Scoring Function Space





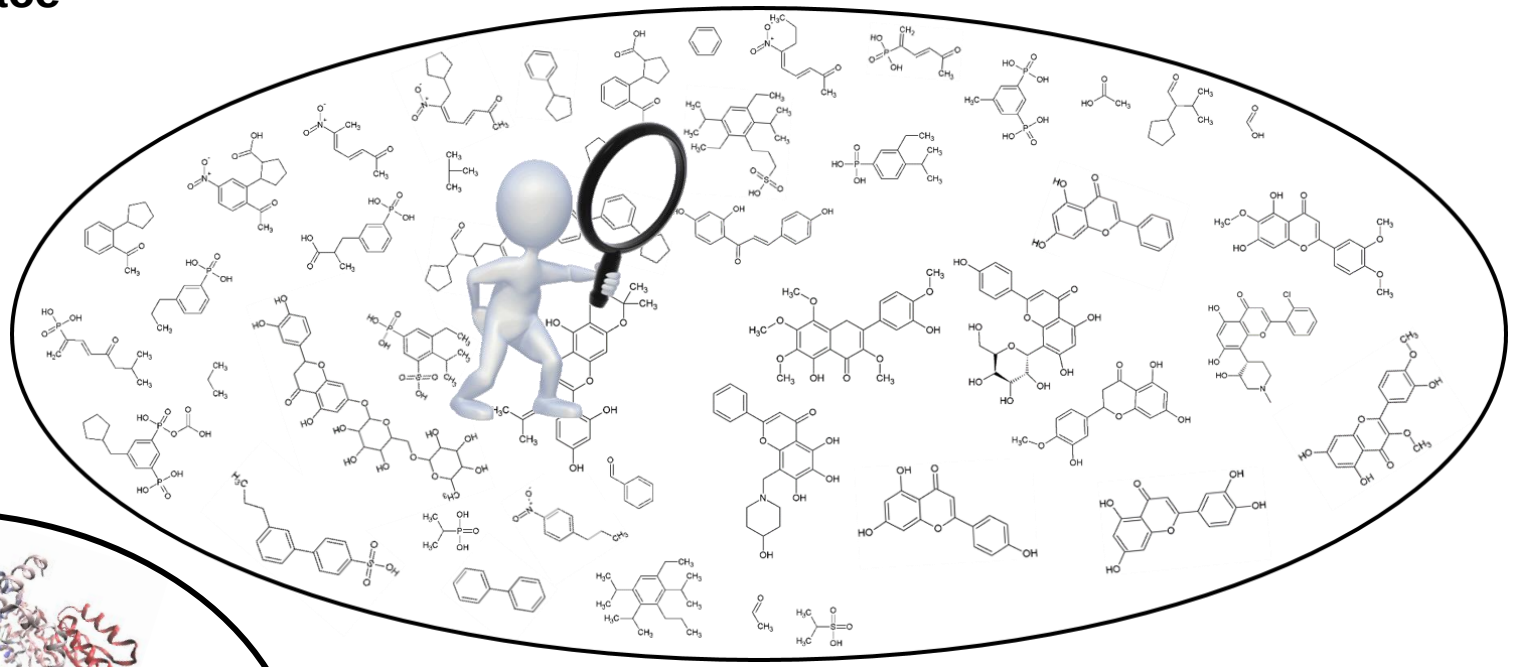
Source: Heck GS, Pintro VO, Pereira RR, de Ávila MB, Levin NMB, de Azevedo WF. Supervised Machine Learning Methods Applied to Predict Ligand- Binding Affinity. Curr Med Chem. 2017; 24(23):2459-2470.

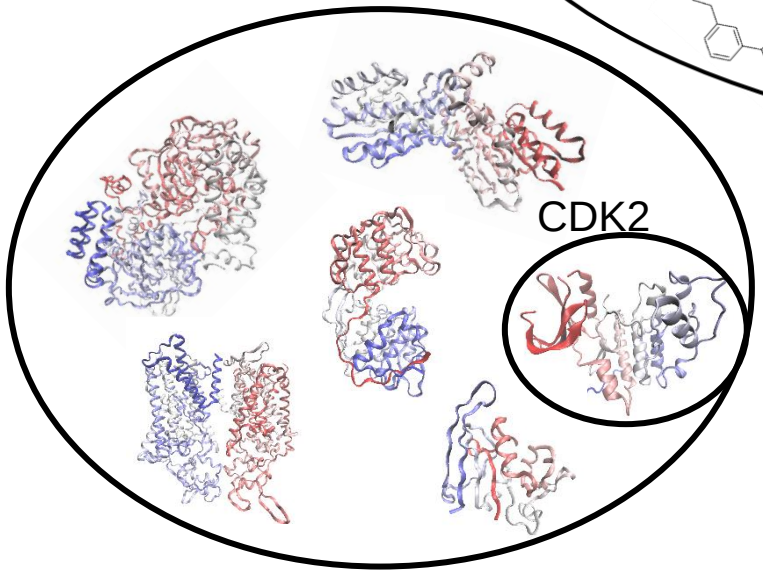
Scoring Function Space



Source: Heck GS, Pintro VO, Pereira RR, de Ávila MB, Levin NMB, de Azevedo WF. Supervised Machine Learning Methods Applied to Predict Ligand- Binding Affinity. Curr Med Chem. 2017; 24(23):2459-2470.

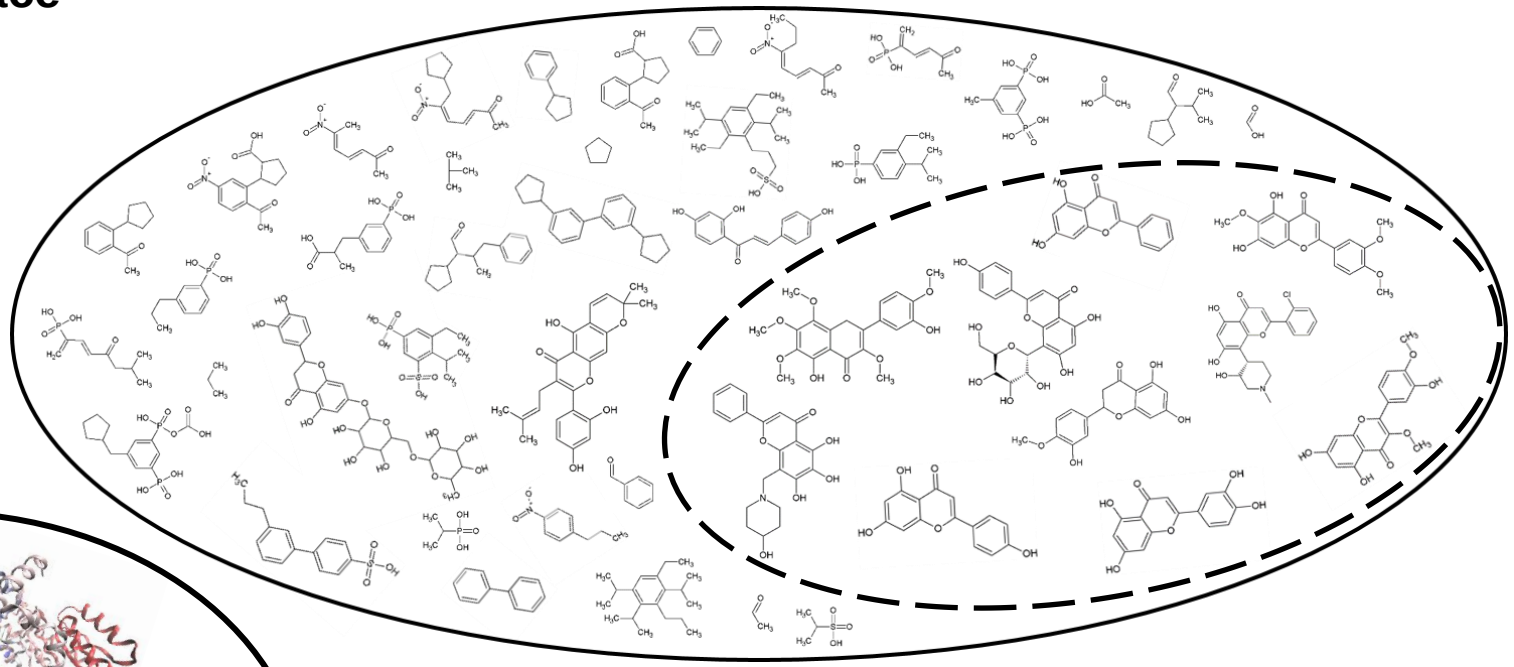
Scoring Function Space

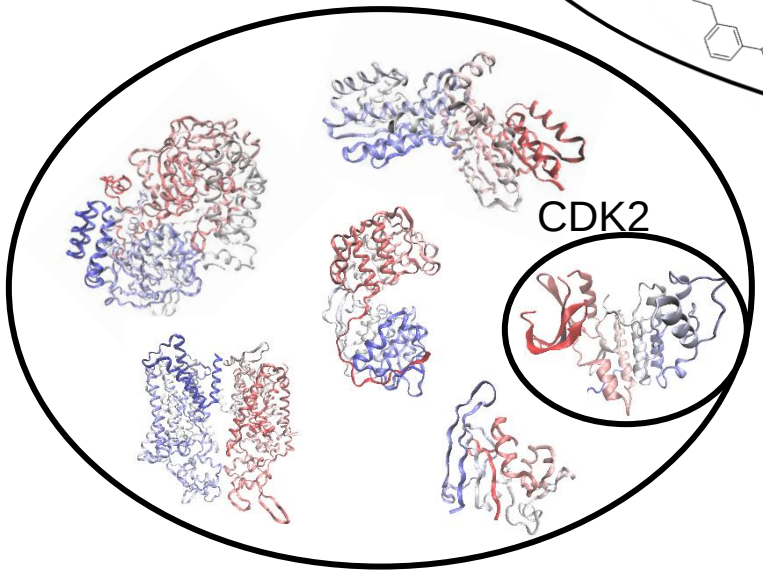




Source: Heck GS, Pintro VO, Pereira RR, de Ávila MB, Levin NMB, de Azevedo WF. Supervised Machine Learning Methods Applied to Predict Ligand- Binding Affinity. Curr Med Chem. 2017; 24(23):2459-2470.

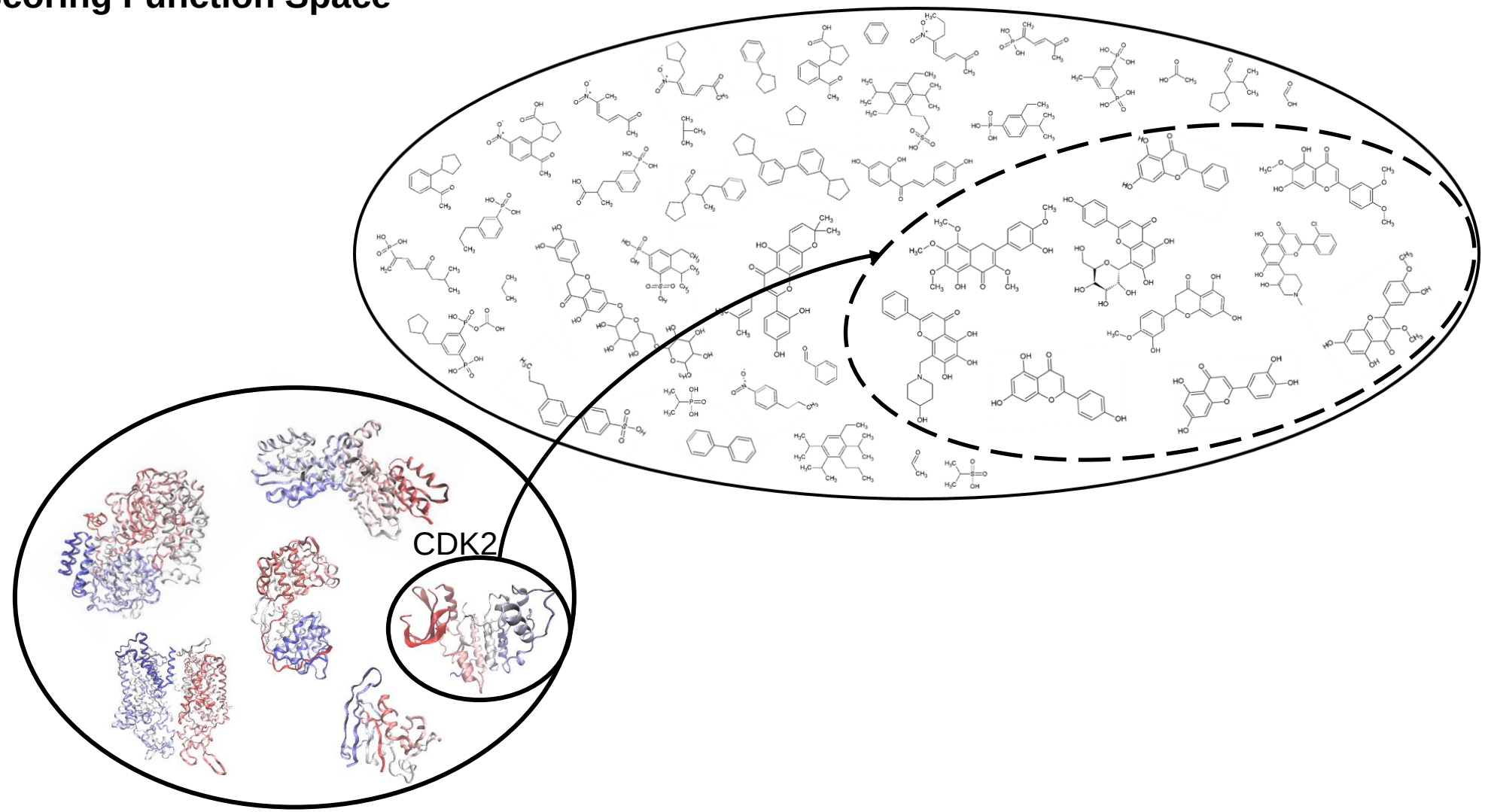
Scoring Function Space





Source: Heck GS, Pintro VO, Pereira RR, de Ávila MB, Levin NMB, de Azevedo WF. Supervised Machine Learning Methods Applied to Predict Ligand- Binding Affinity. Curr Med Chem. 2017; 24(23):2459-2470.

Scoring Function Space



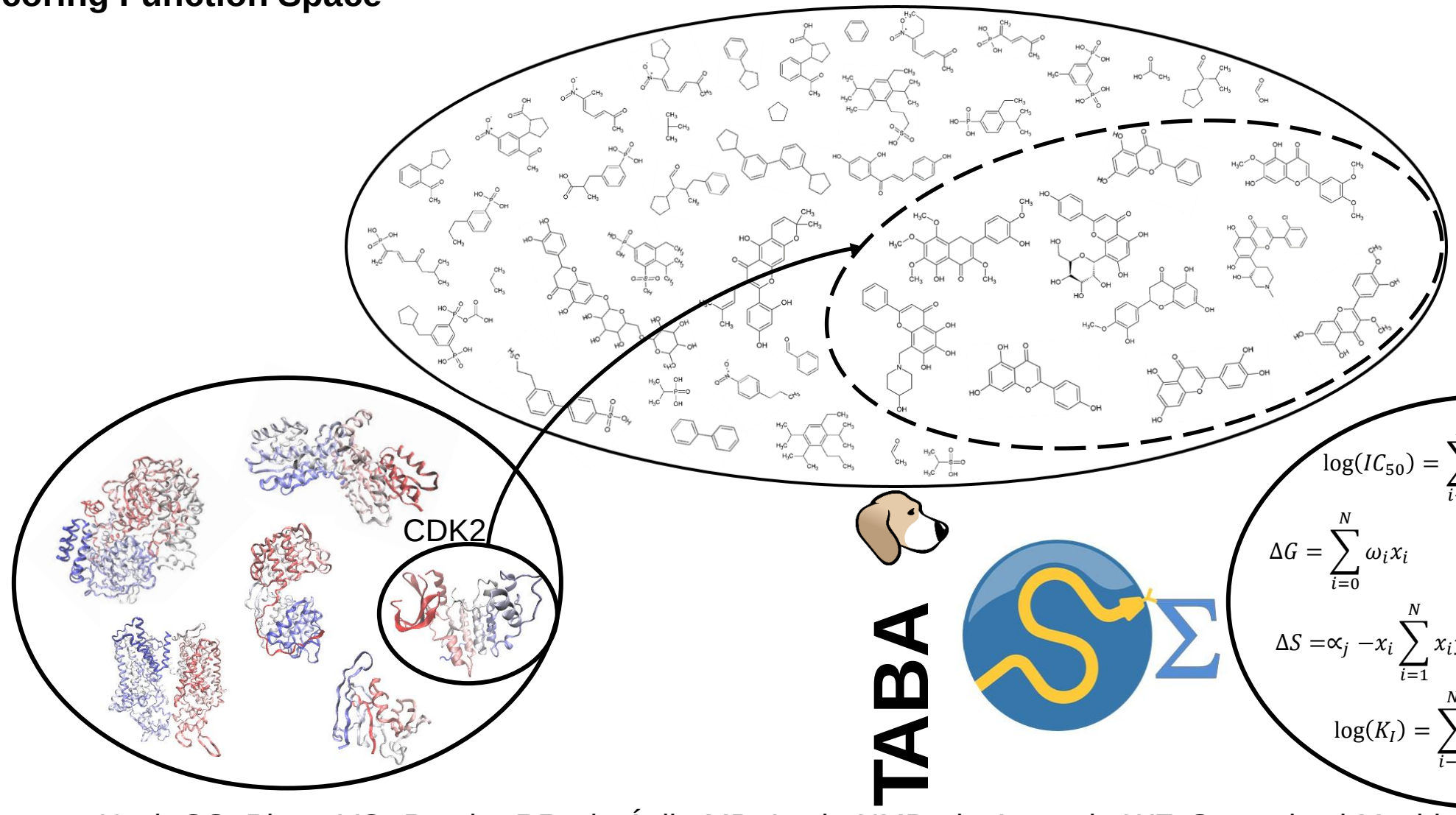
Binding MOAD Mother of All Databases

PDBbind

BindingDB

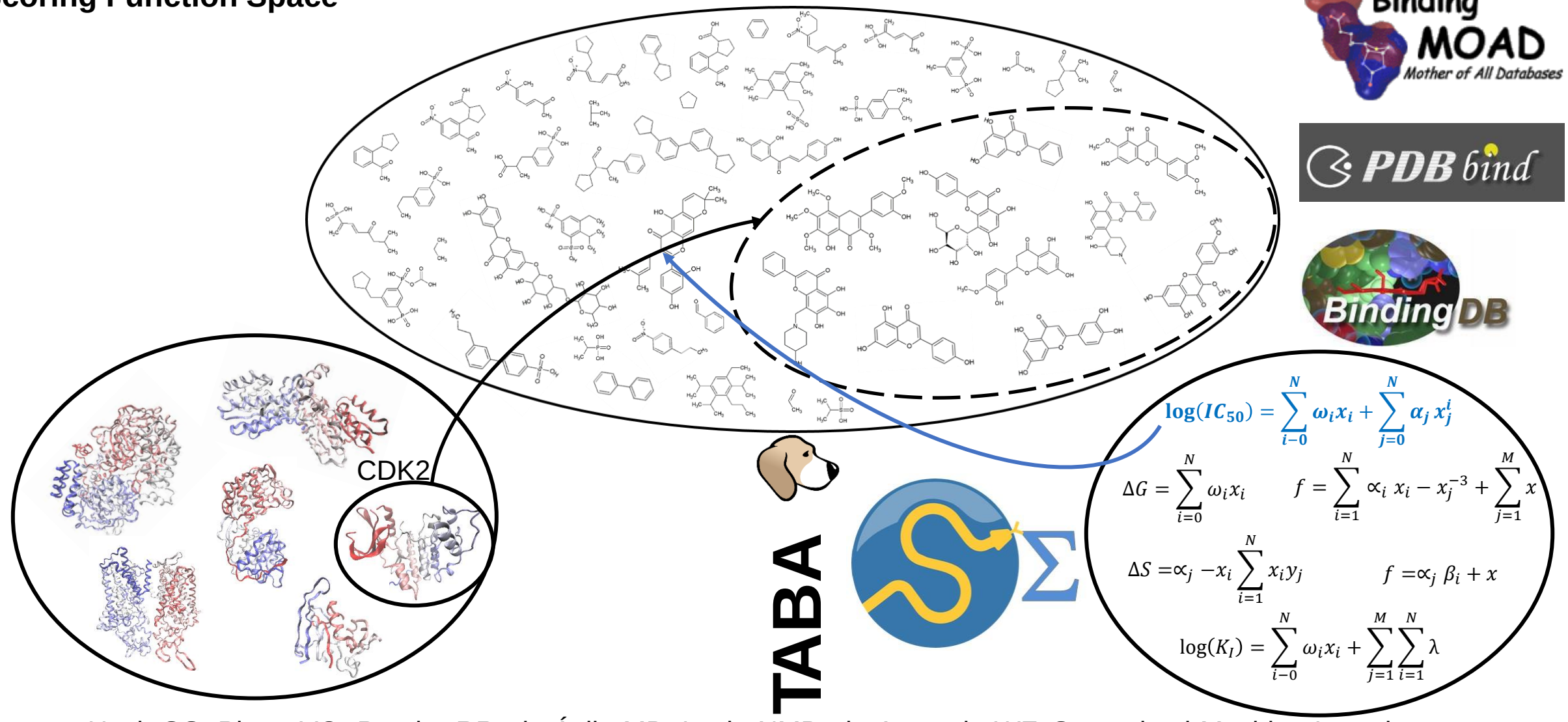
Source: Heck GS, Pintro VO, Pereira RR, de Ávila MB, Levin NMB, de Azevedo WF. Supervised Machine Learning Methods Applied to Predict Ligand- Binding Affinity. Curr Med Chem. 2017; 24(23):2459-2470.

Scoring Function Space



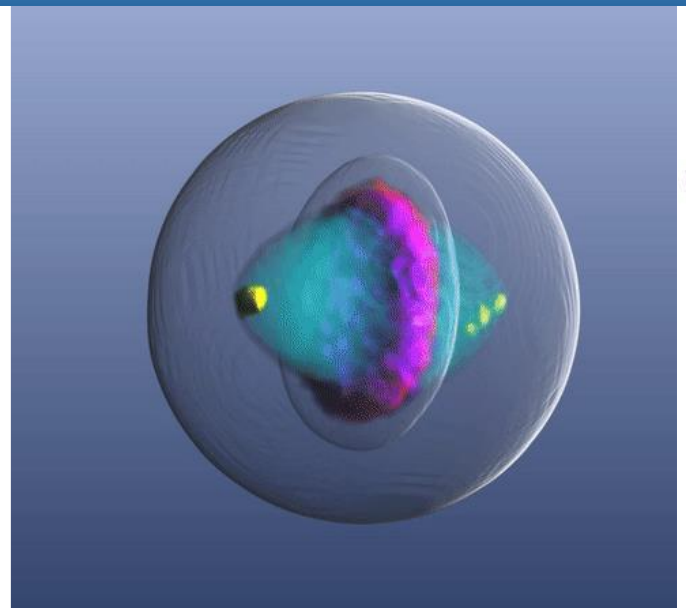
Source: Heck GS, Pintro VO, Pereira RR, de Ávila MB, Levin NMB, de Azevedo WF. Supervised Machine Learning Methods Applied to Predict Ligand- Binding Affinity. Curr Med Chem. 2017; 24(23):2459-2470.

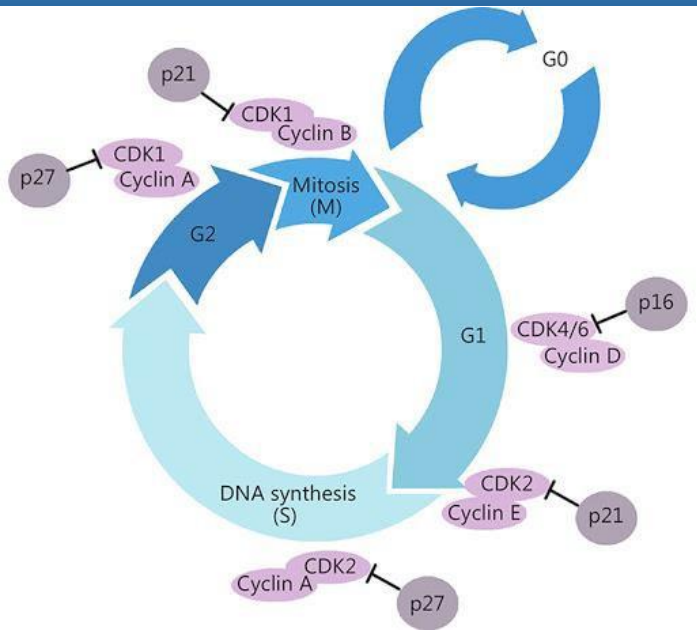
Scoring Function Space

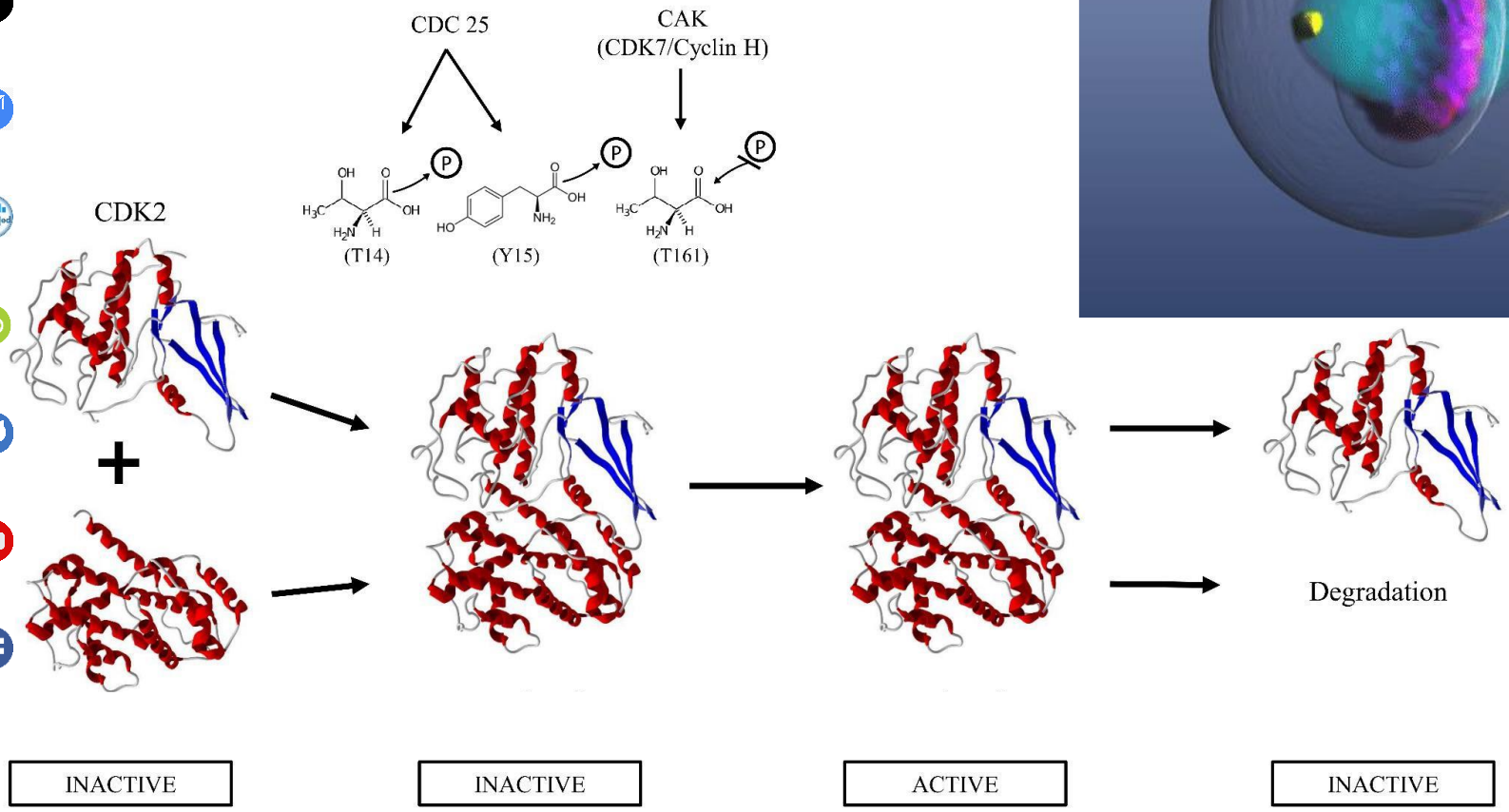


Source: Heck GS, Pintro VO, Pereira RR, de Ávila MB, Levin NMB, de Azevedo WF. Supervised Machine Learning Methods Applied to Predict Ligand- Binding Affinity. Curr Med Chem. 2017; 24(23):2459-2470.

Cyclin-Dependent Kinase 2



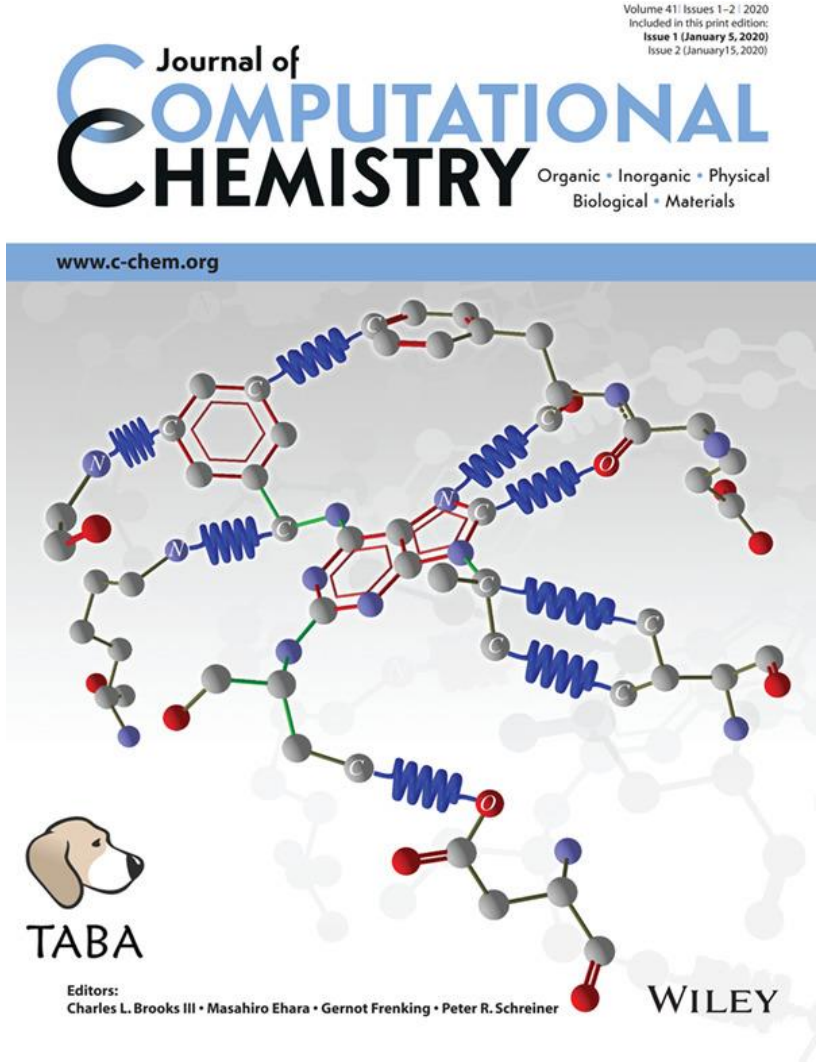


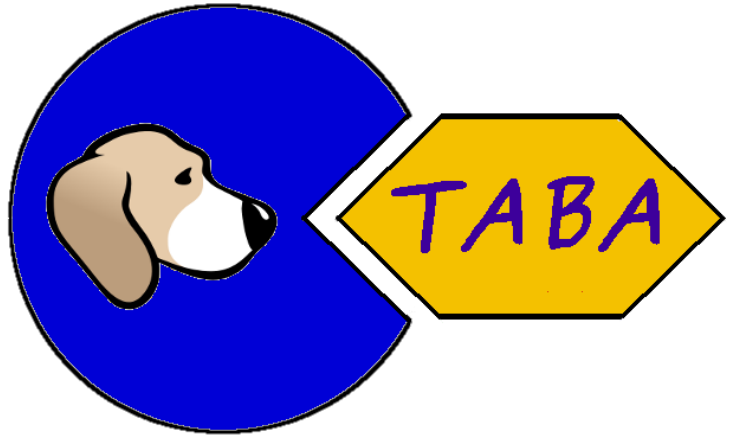


RCSB PDB PROTEIN DATA BANK

Source: de Azevedo WF. Protein-Ligand Interactions: High-Resolution Structures of CDK2. Curr Drug Targets. 2022;23(5):438-440.

Taba-Tool to Analyze the Binding Affinity





python powered

NumPy

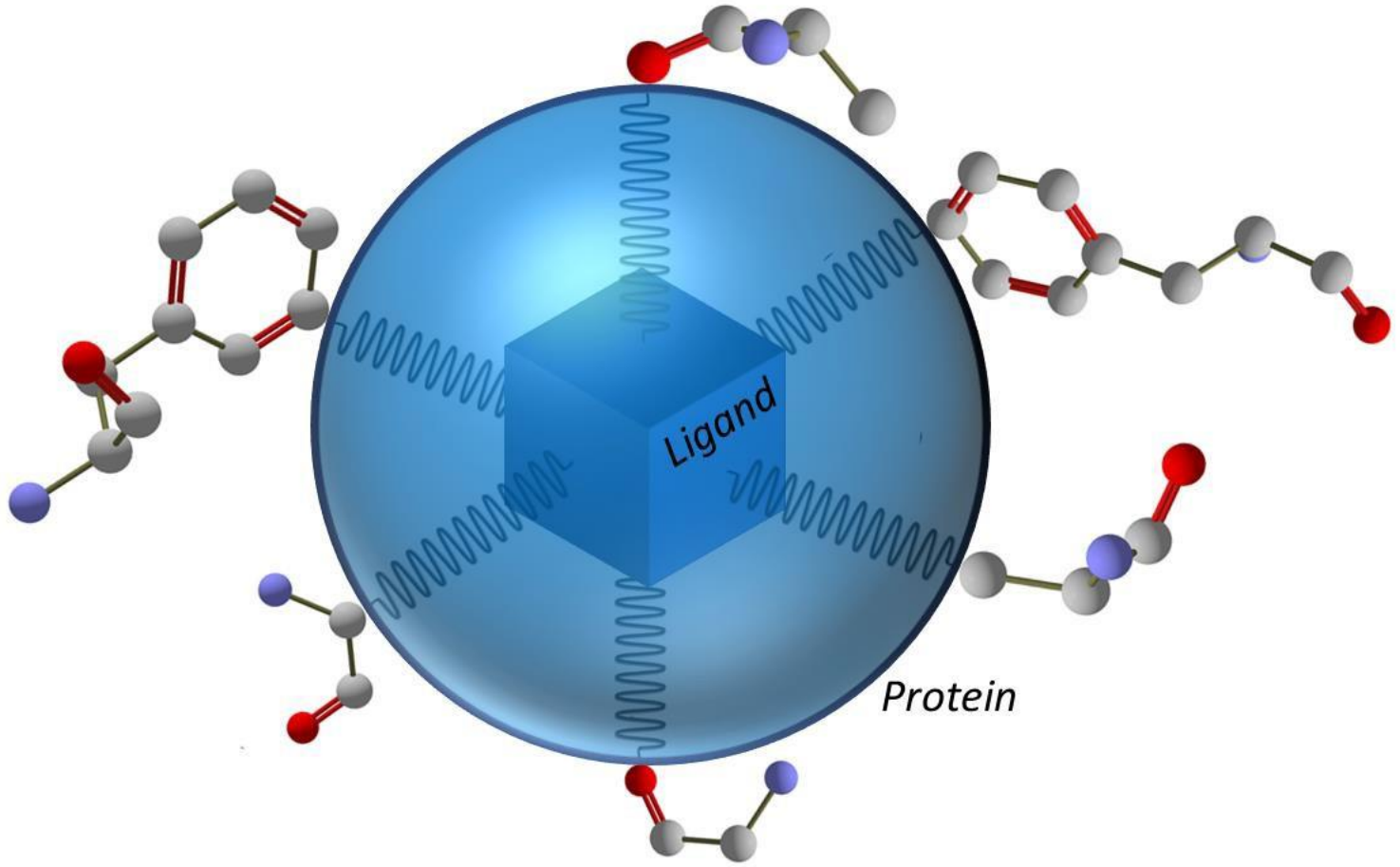
SciPy.org Sponsored By ENTHOUGHT

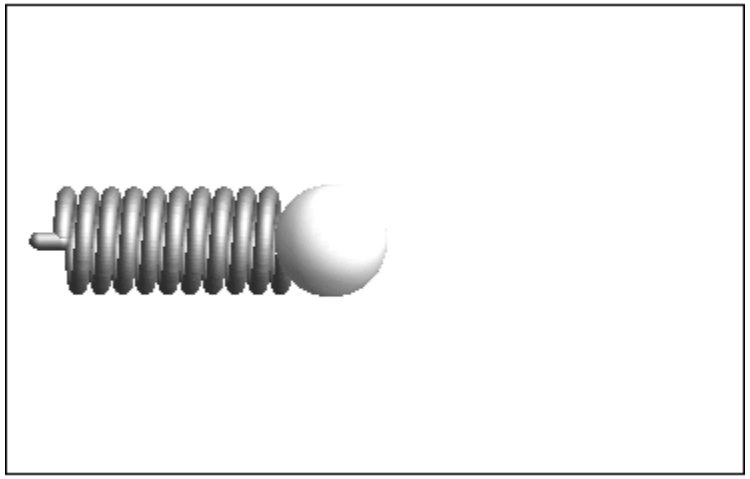
scikit learn

Available at: https://github.com/azevedolab

Source: da Silva AD, Bitencourt-Ferreira G, de Azevedo WF Jr. Taba: A Tool to Analyze the Binding Affinity. J Comput Chem. 2020; 41(1):69-73.

Taba-Tool to Analyze the Binding Affinity





Source: da Silva AD, Bitencourt-Ferreira G, de Azevedo WF Jr. Taba: A Tool to Analyze the Binding Affinity. J Comput Chem. 2020; 41(1):69-73.

Taba-Tool to Analyze the Binding Affinity

Scoring Functions	ρ	p-value1	R^2	p-value2
Free Energy[a]	-0.133	0.7324	0.204	0.2227
Final Intermolecular Energy[a]	0.133	0.7324	0.204	0.2228
vdW+Hbond+desolv Energy[a]	0.133	0.7324	0.204	0.2228
Electrostatic Energy[a]	0.533	0.1392	0.376	0.0789
Final Total Internal Energy[a]	-0.133	0.7324	0.089	0.4365
Torsional Free Energy[a]	0.068	0.8630	0.000	0.9792
Plants Score[b]	0.183	0.6368	0.001	0.9348
MolDock Score[b]	0.217	0.5755	0.010	0.7950
Rerank Score[b]	0.333	0.3807	0.007	0.8336
Interaction Score[b]	0.367	0.3317	0.013	0.7698
Protein Score[b]	0.367	0.3317	0.025	0.6839
Water Score[b]	-0.569	0.1098	0.395	0.0699
Internal Score[b]	0.033	0.9322	0.001	0.9369
Electrostatic Score[b]	0.548	0.1269	0.204	0.2218
Electrostatic Long Score[b]	-0.548	0.1269	0.204	0.2218
H-Bond Score[b]	0.650	0.0581	0.512	0.0301
Ligand Efficiency 1 Score[b]	0.150	0.7001	0.024	0.6935
Ligand Efficiency 3 Score[b]	0.283	0.4600	0.023	0.6968
Affinity Score[c]	-0.067	0.8647	0.117	0.3669
Gauss1 Score[c]	-0.367	0.3317	0.120	0.3603
Gauss2 Score[c]	-0.283	0.4600	0.018	0.7297
Repulsion Score[c]	-0.700	0.0358	0.240	0.1804
Hydrophobic Score[c]	0.100	0.7980	0.002	0.9157
Hydrogen Score[c]	-0.583	0.0992	0.340	0.0993
Taba (3 variables, d≤ 4.5 Å)	0.783	0.01252	0.794	0.0107

Predictive performance of scoring functions (test set). [a]AutoDock 4, [b]Molegro Virtual Docker (MVD), [c]AutoDock Vina. p-value1 and p-value2 are related to ρ and R, respectively.

SAnDReS-Statistical Analysis of Docking Results and Scoring functions

- Sixty-four regression methods
- AutoDock Vina 1.2.3 force field
- Open source

Available at: https://github.com/azevedolab

SAnDReS-Statistical Analysis of Docking Results and Scoring functions

Method	r	p-value(r)	r2	rho	p-value(rho)	MSE	RMSE
GaussianProcessRegressorCV	0.419643	0.300663	0.1761	0.512348	0.194223	2.19599	1.48189
GaussianProcessRegressor	0.332622	0.420807	0.110638	0.439155	0.276327	3.52633	1.87785
DecisionTreeRegressorCV	0.225384	0.591498	0.0507979	0.273297	0.512512	2.48651	1.57687
ExtraTreesRegressorCV	0.224615	0.592798	0.0504519	0.414758	0.306912	2.44803	1.56462
MLPRegressorCV	0.210381	0.61702	0.0442603	0.317168	0.443989	3.48123	1.8658
RANSACRegressor	0.196385	0.641136	0.038567	0.365963	0.372625	3.06544	1.75084
PassiveAggressiveRegressor	0.194978	0.643576	0.0380165	0.365963	0.372625	1.65285	1.28563
SVRCV	0.192942	0.647111	0.0372267	0.365963	0.372625	2.06293	1.43629
MLPRegressor	0.1926	0.647706	0.0370949	0.121988	0.773532	2.11687	1.45495
SVR	0.192141	0.648505	0.036918	0.365963	0.372625	1.80235	1.34251
LinearSVRCV	0.191566	0.649504	0.0366977	0.463553	0.247323	1.68465	1.29794
NuSVRCV	0.191518	0.649588	0.0366792	0.463553	0.247323	1.94074	1.3931
RANSACRegressorCV	0.191485	0.649645	0.0366665	0.317168	0.443989	2.70793	1.64558
TheilSenRegressor	0.190028	0.652182	0.0361107	0.463553	0.247323	1.54587	1.24333
TweedieRegressorCV	0.189321	0.653414	0.0358425	0.463553	0.247323	1.4201	1.19168
PassiveAggressiveRegressorCV	0.189212	0.653604	0.0358012	0.317168	0.443989	2.26334	1.50444
BayesianRidgeCV	0.18763	0.656364	0.035205	0.463553	0.247323	1.44061	1.20025
LarsCV	0.186977	0.657504	0.0349603	0.463553	0.247323	1.5377	1.24004
TweedieRegressor	0.186895	0.657647	0.0349297	0.365963	0.372625	1.30757	1.14349
ExtraTreeRegressor	0.186558	0.658235	0.0348039	0.253185	0.545176	2.99684	1.73114
LinearSVR	0.186361	0.65858	0.0347304	0.463553	0.247323	1.68322	1.29739
TheilSenRegressorCV	0.184789	0.661328	0.0341469	0.365963	0.372625	2.50186	1.58173
BayesianRidge	0.184484	0.66186	0.0340344	0.365963	0.372625	1.31805	1.14806

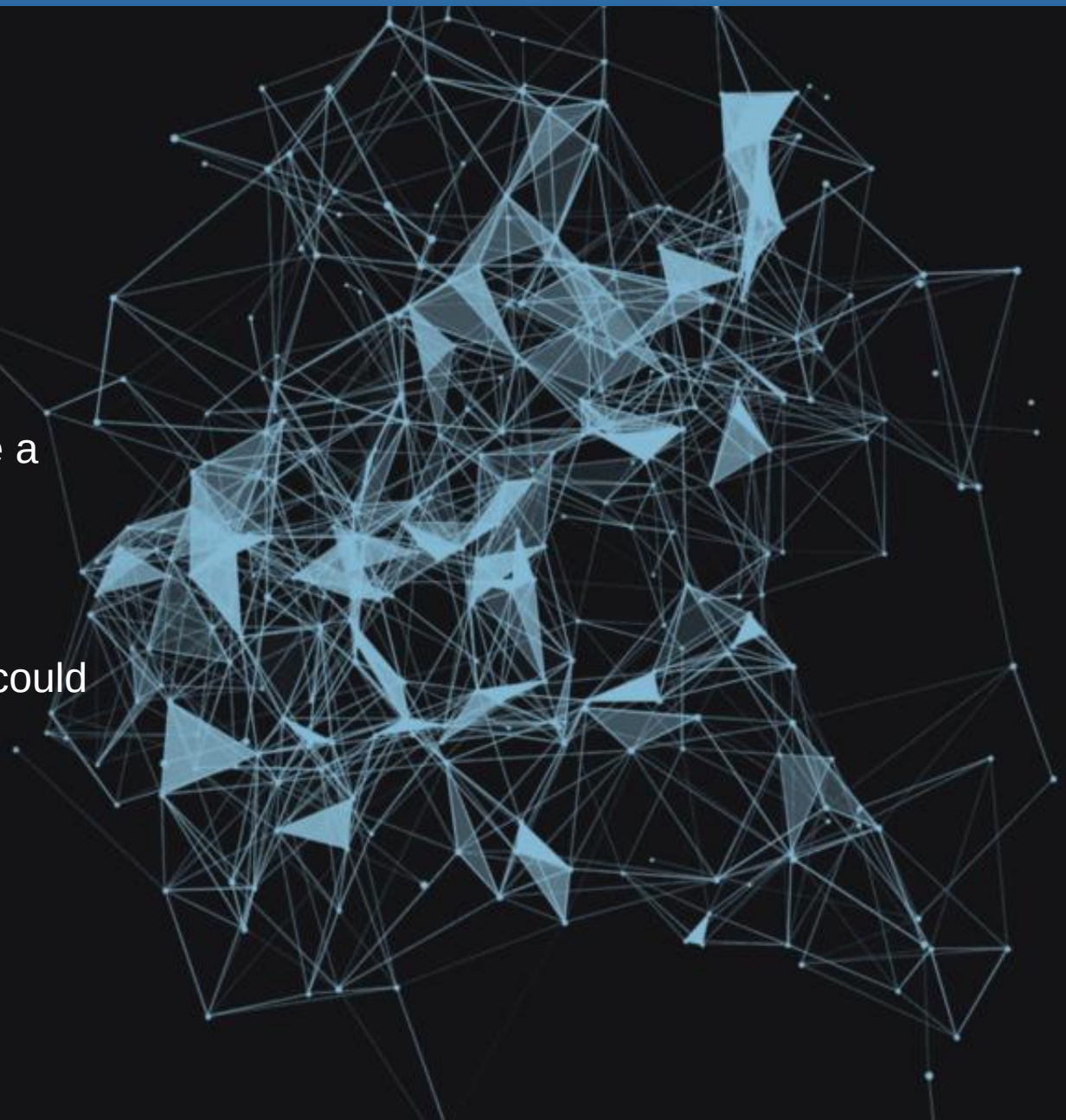
Our machine learning models generalize better than the AutoDock Vina scoring function, which has an **r^2 = 0.00909986.**

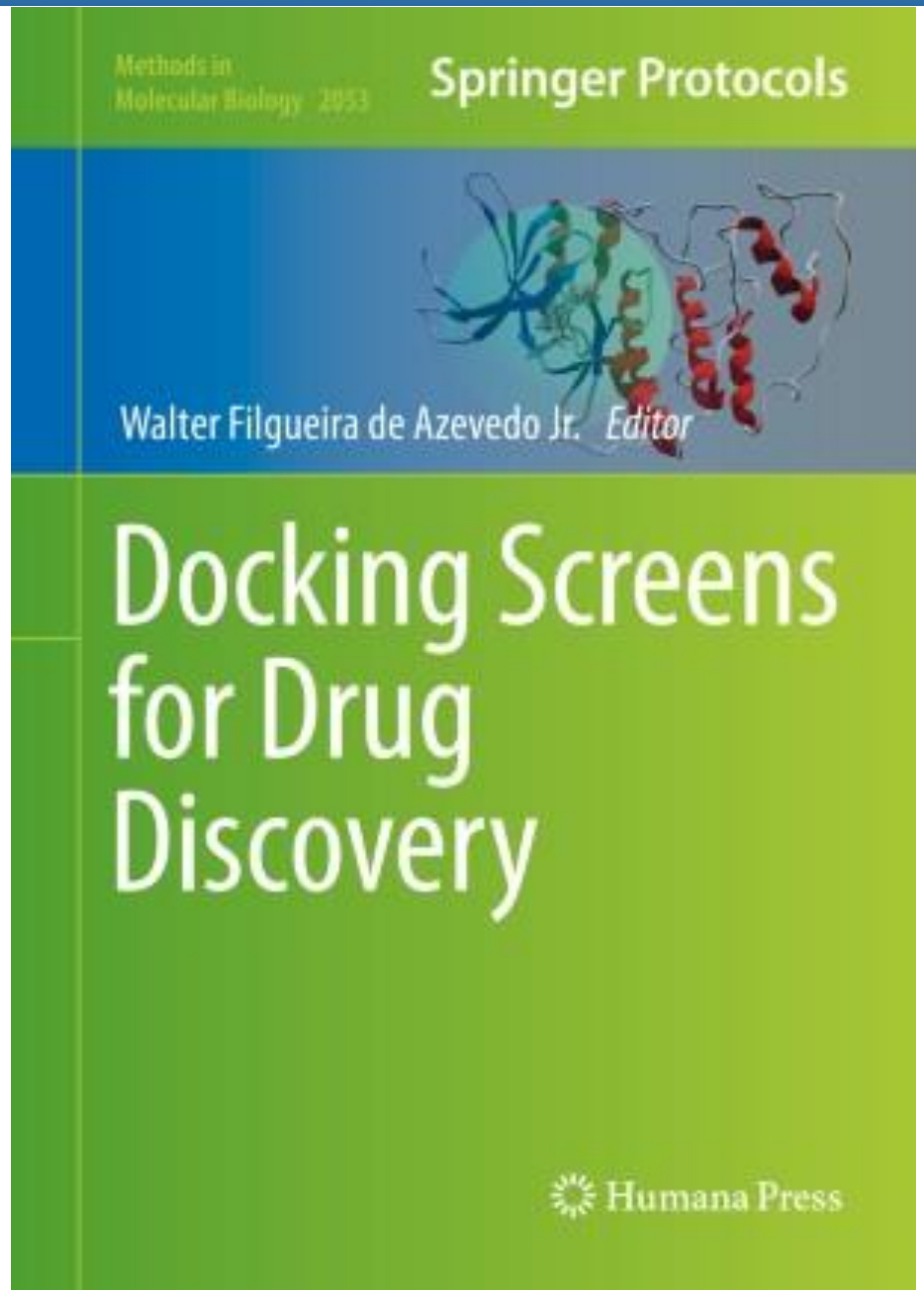
Final Remarks

- Taba and SAnDReS allow us to explore the Scoring Function Space with efficiency;

- We may use Taba and SAnDReS to generate a database of scoring functions – SF database;

- We believe that the Scoring Function Space could be a game-changer for drug development.




Methods in Molecular Biology 2053
Springer Protocols
Walter Filgueira de Azevedo Jr. Editor
Docking Screens for Drug Discovery
Humana Press

Young Scientists flash presentations
Exploring the Scoring Function Space for Structure-Based Drug Design

Wednesday May 24, 2020
6:20 PM

Gabriela Bitencourt-Ferreira
(Ph.D. Student) (H-index: 12)
azevedolab.net

gabriela@azevedolab.net

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
Porto Alegre/RS
Brazil

Acknowledgements

Dr. Olga Tarasova - Institute of Biomedical Chemistry, Moscow-Russia
Prof. Dr. Vladimir Poroikov - Institute of Biomedical Chemistry, Moscow-Russia

Organizing Committee of XXVIII Symposium on Bioinformatics and Computer-Aided Drug Discovery

Acknowledgements

Ph.D Students

Maurício Boff de Ávila

Silvana da Cunha Russo

Gabriela Bitencourt-Ferreira

Master Student

Camila Rizzotto

Specialization Student

Amauri Duarte da Silva

Undergraduate Student

Martina Veit Acost

Acknowledgements

Thank you!

Acknowledgements

CNPq

Capes

Acknowledgements

Harnessing Machine Learning for Drug Development

Dr. Walter F. de Azevedo, Jr.
azevedolab.net

walter@azevedolab.net

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)
Porto Alegre/RS
Brazil

https://publons.com/researcher/1890214/walter-f-de-azevedo-jr/

Big news! Your Publons™ profile is moving to the Web of Science™. Click here to read more.

publons BROWSE COMMUNITY FAQS LOG IN REGISTER WEB OF SCIENCE

Home ▸ Researchers ▸ Walter F de Azevedo, Jr.

Walter F de Azevedo, Jr.

"Walter Azevedo Jr."

Show more

Pontifícia Universidade Católica do Rio Grande do Sul

Web of Science ResearcherID
N-5585-2018

PUBLICATIONS	TOTAL TIMES CITED	H-INDEX	VERIFIED REVIEWS	VERIFIED EDITOR RECORDS
204	6651	45	271	37

- Summary
- Metrics
- Publications
- Peer review

Research Fields

BIOCHEMISTRY & MOLECULAR BIOLOGY · BIOINSPIRED COMPUTING · BIOPHYSICS · COMPUTATIONAL SYSTEMS BIOLOGY · COMPUTER SCIENCE · CRYSTALLOGRAPHY · MACHINE LEARNING · MATHEMATICAL & COMPUTATIONAL BIOLOGY · MOLECULAR DOCKING · PROTEIN-LIGAND INTERACTIONS

+ VIEW ALL INSTITUTIONS & AWARDS