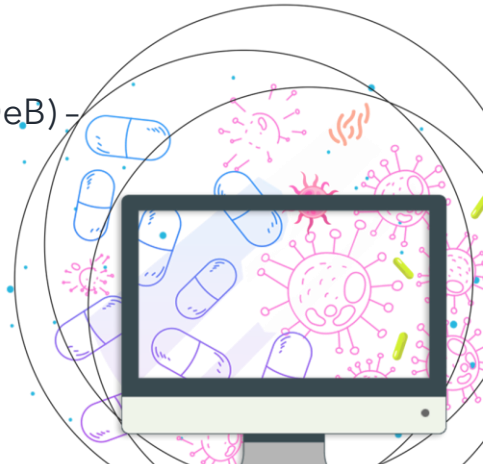
XXXI Symposium on Bioinformatics and Computer-Aided Drug Discovery (BCADD-2025)

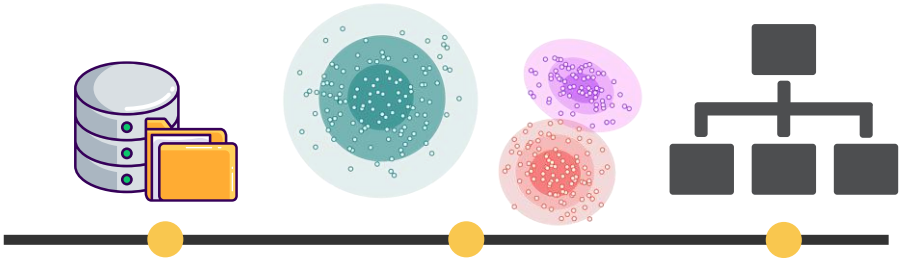
Emerging Challenges and Opportunities for In Silico Drug Discovery

COMBINING MACHINE LEARNING AND STRUCTURE-BASED APPROACHES FOR THE EFFICIENT IDENTIFICATION OF NOVEL BIOACTIVE SCAFFOLDS

Alan Talevi

Laboratory of Bioactive Compound Research and Development (LIDeB) – University of La Plata / Boolzi

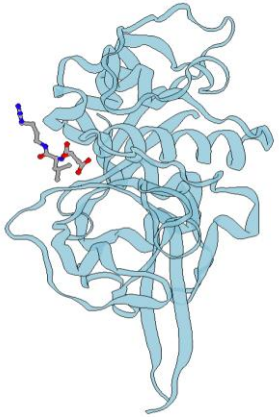




Data set compilation & curation

Clustering

Sampling



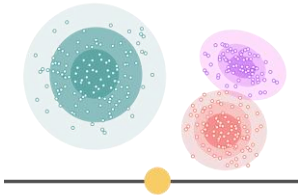
Supervised ML

In silico screening

Hit selection

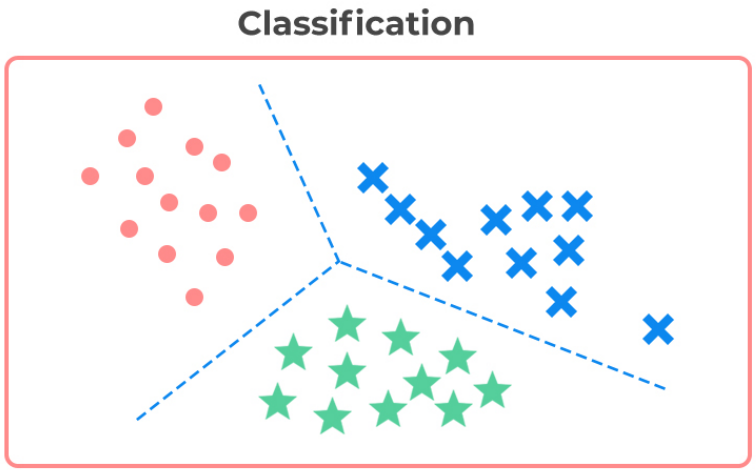
Structure-based approaches

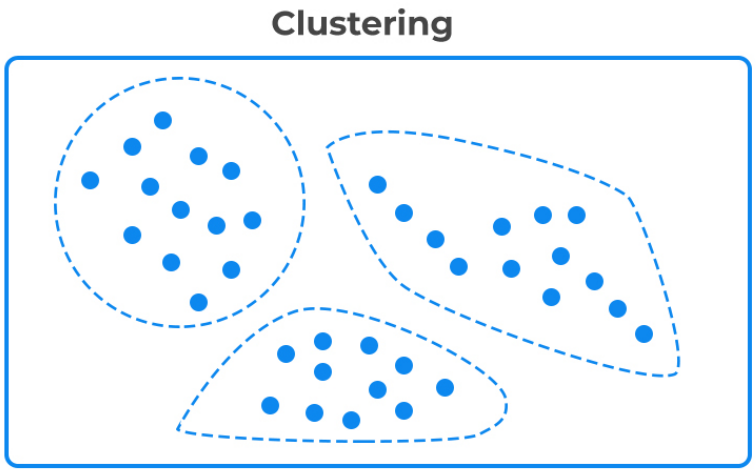
In vitro assays



Clustering

machine learning



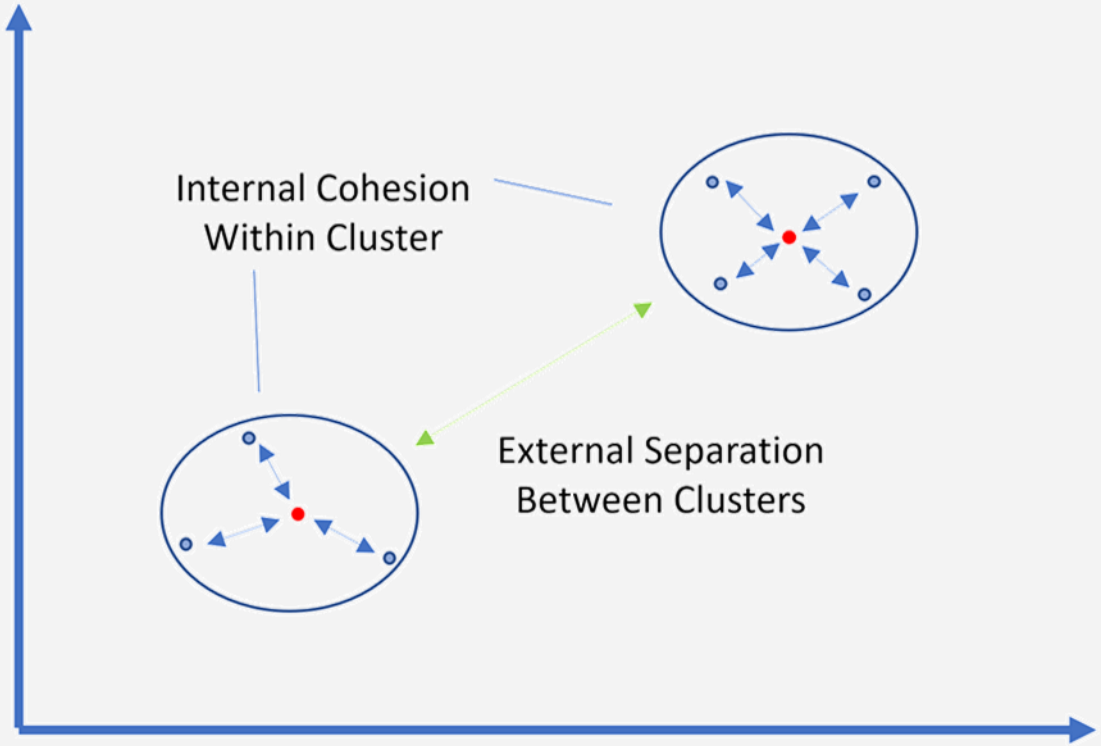


In comparative terms, unsupervised machine learning is far less frequently used than supervised machine learning for structuring small-molecule data.

How can the validity of a proposed clustering be assessed?

- Internal and external validation
- Stability / replicability
- Relative validation

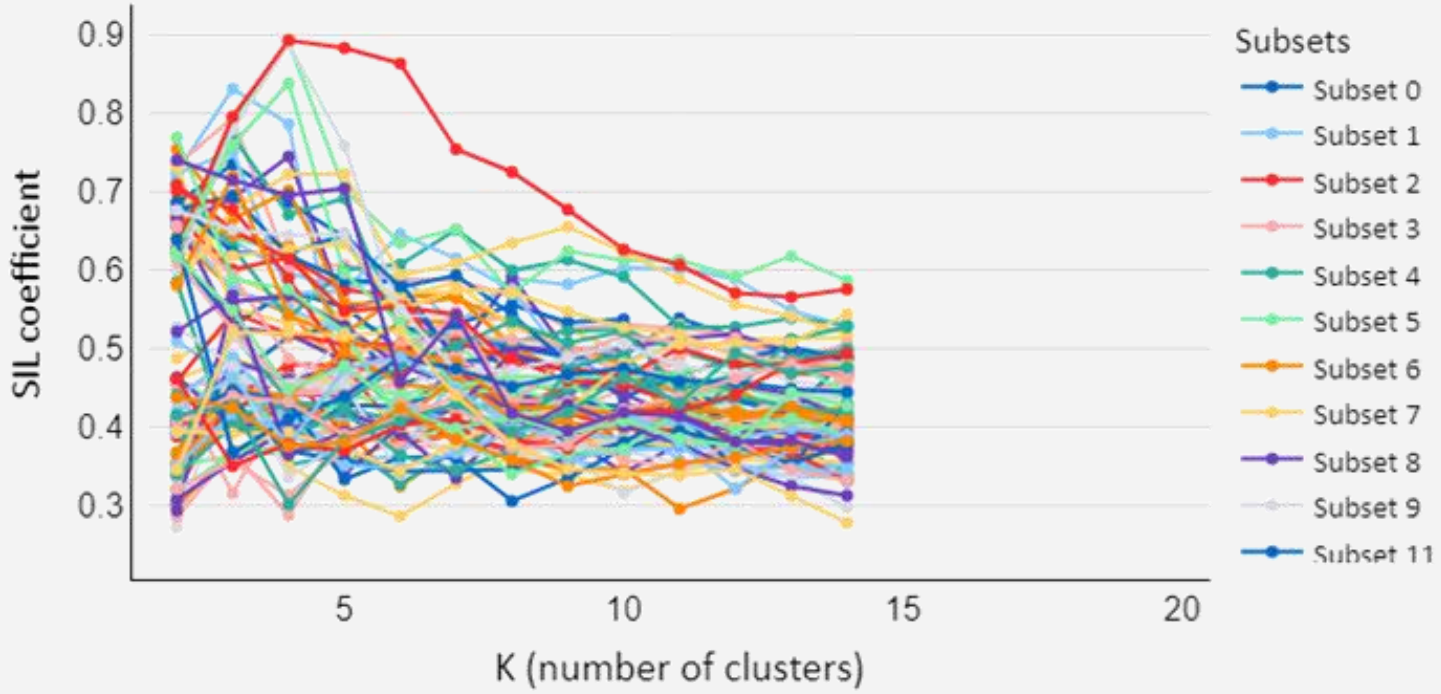
Internal validation



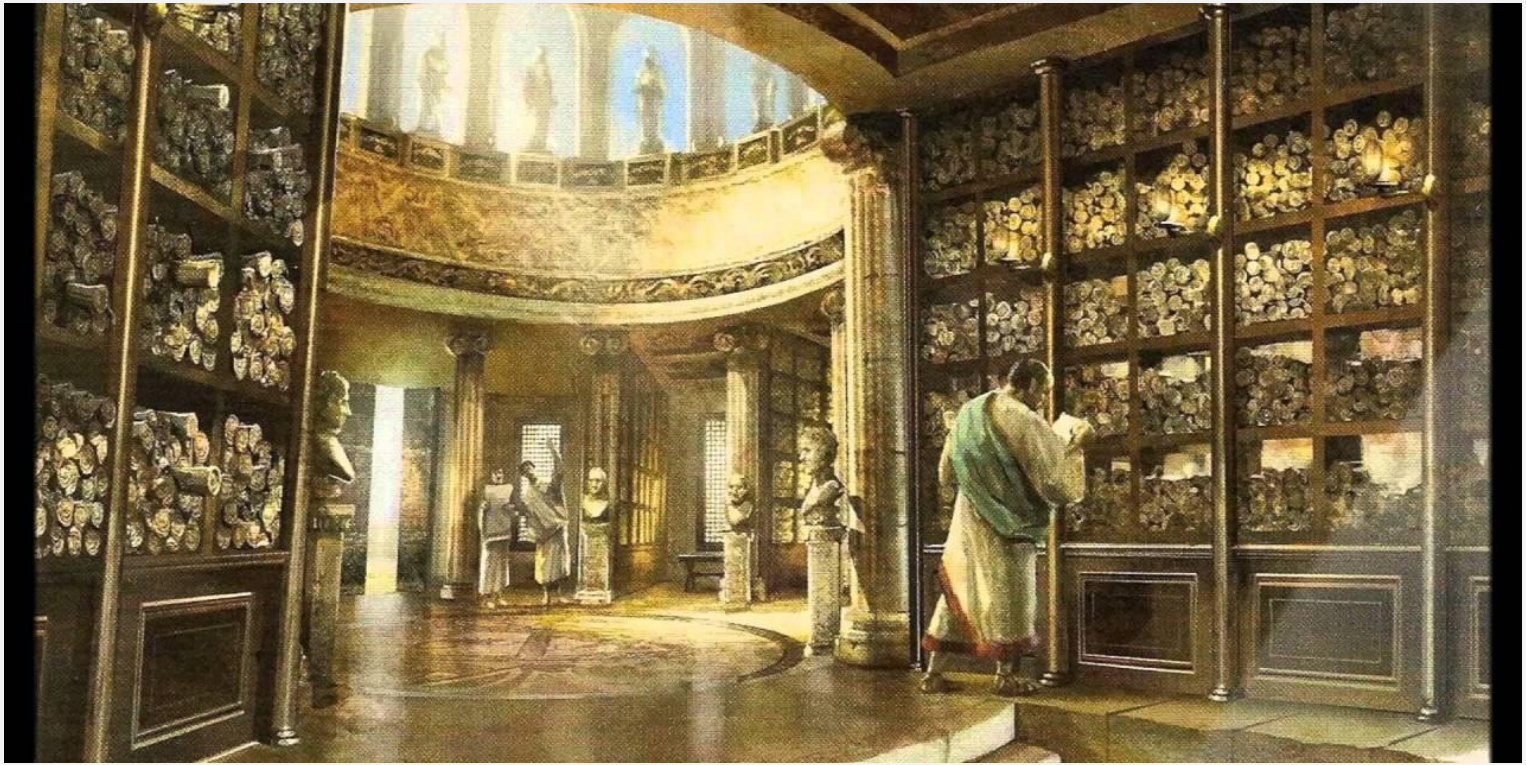
$$S(i) = \frac{b(i) - a(i)}{\max\{\hat{a}(i), b(i)\}}$$

a(i) is the average distance between element i and all the other elements within its own cluster, while b(i) is the average distance between i and the elements of the nearest cluster (excluding its own)

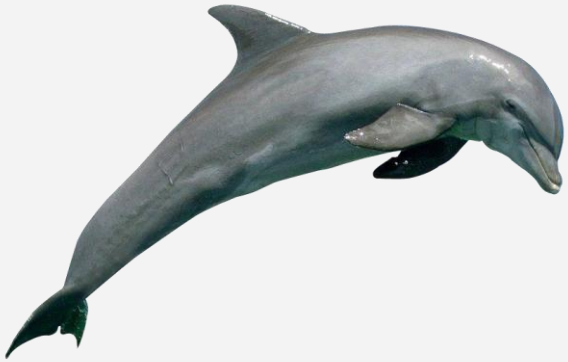
Relative validation




Literary Fiction
Thriller & Mystery
New and Notable
RUTH OZEKI
LOUISE PENNY
THE PLOT
LEVIATHAN WAKES
THE TURNOUT
BATH HAUS
A SLOW FIRE BURNING
PAULA HAWKINS



0	0	0	1	1	1



1	1	1	1	1	1

Caudal fin
Caudal fin
Dorsal fin
Dorsal fin
Dorsal fin
Lungs
Pectoral fin
Pectoral fin
Mammary glands
Diaphragm

1	1	1	0	0	0

Conclusion 1

The ability of an algorithm to cluster, classify, etc. (and the results it produces) will fundamentally depend on the features (feature space) used to perform these tasks.

Corollary: there would be no such thing as a “natural” clustering (?)

Fiction

Science fiction

A B C D...

Horror

A B C D...

Noir

A B C D...

Historical

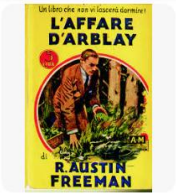
A B C D...

Conclusion 2

Obtaining subclusters is sometimes useful

[especially when working with large sets of elements to be clustered]

Google libri gialli



Wikipedia
Il Giallo Mondadori - …


LFB.it
FFF - Mondadori: I libri gialli,…

ilLibraio.it
Perché i romanzi gialli…

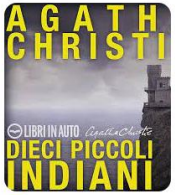
Focus Junior
Libri gialli per ragazzi …

Roberto Tartaglia
I libri gialli da leggere …

Marie Claire
Libri gialli: classici e nuovi da leggere | Marie …

Thriller Cafe
Libri gialli con Repubbl…

Consigli.it
Scopri i 5 libri gialli imperdibili d…



Audible
Migliori libri gialli: le …

Fanpage
I migliori libri gialli del 2021 da leggere ass…

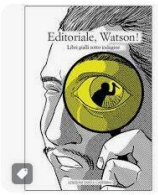
ELLE
10 nuovi libri gialli …

d11gmip42rcud8.cl…
Qué es el cine Gial…

IBS · Disponibile
Il grande libro dei …

India Nepal Viaggi
L'India misteriosa in 10 libri gialli …

Libri News



Edizioni S… · Disponibile

Esquire

Focus Junior

De Agostini Libri

Biblioteca Salaborsa

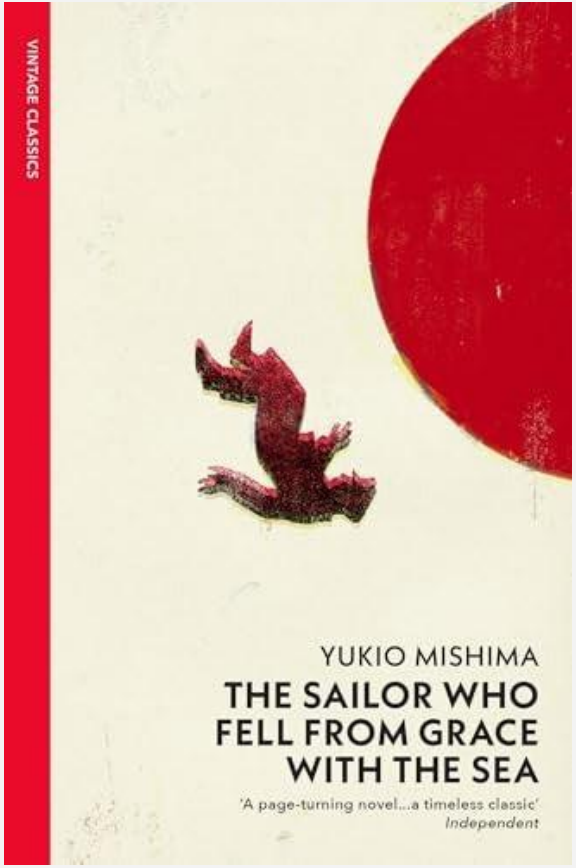
Esquire

Conclusion 3

From time to time, unexpected features may yield useful clusterings

[even if we fail to fully elucidate the underlying logic]


MURAKAMI
1Q84
BOOK THREE
VINTAGE


VINTAGE CLASSICS
YUKIO MISHIMA
THE SAILOR WHO FELL FROM GRACE WITH THE SEA
'A page-turning novel...a timeless classic'
Independent

Conclusion 4

When two objects (e.g., molecules) consistently group together across multiple feature subspaces, their co-localization may carry special significance.

Summarizing

- Different feature spaces may provide different results
- Iterative clustering may be productive to structure big data
- Unsuspected features may sometimes be useful
- The consistency of a data structure across different feature subspaces or algorithms could be meaningful.

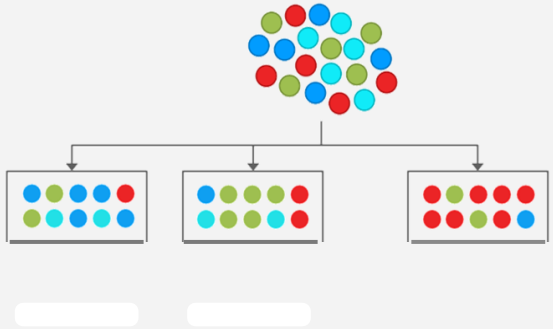
JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

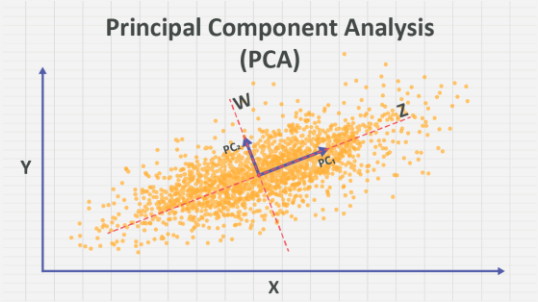
iRaPCA and SOMoC: Development and Validation of Web Applications for New Approaches for the Clustering of Small Molecules

Denis N. Prada Gori, Manuel A. Llanos, Carolina L. Bellera, Alan Talevi,* and Lucas N. Alberca*

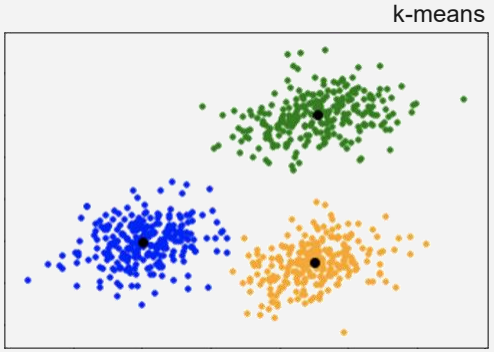
Cite This: https://doi.org/10.1021/acs.jcim.2c00265



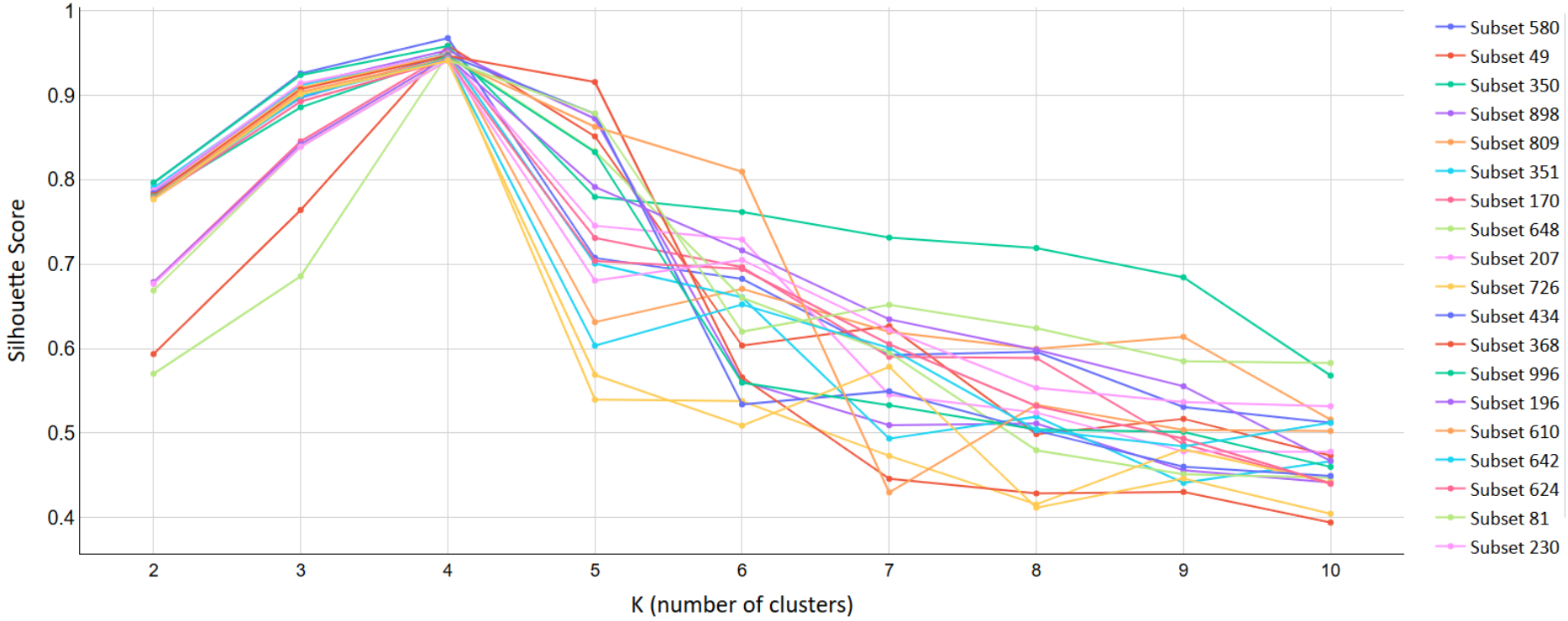
Feature bagging



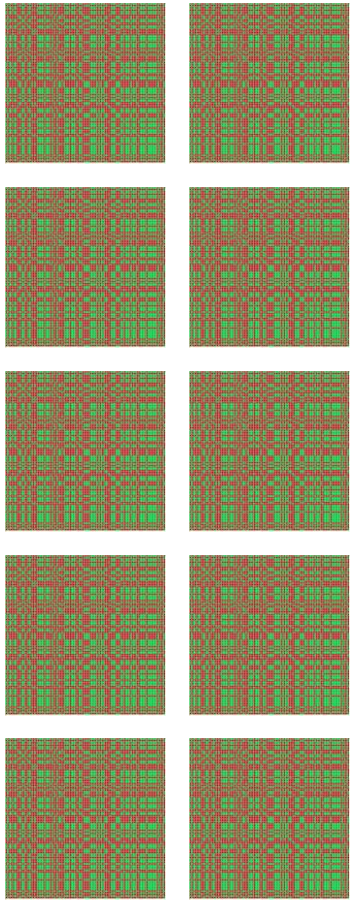
Dimensionality reduction

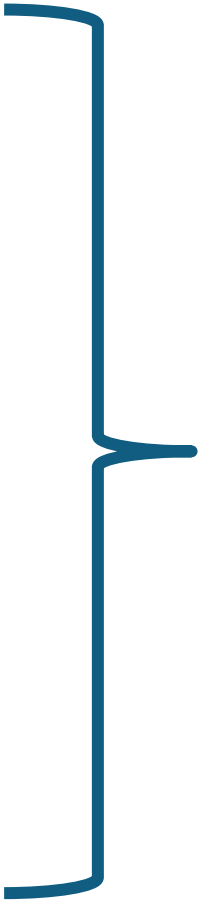


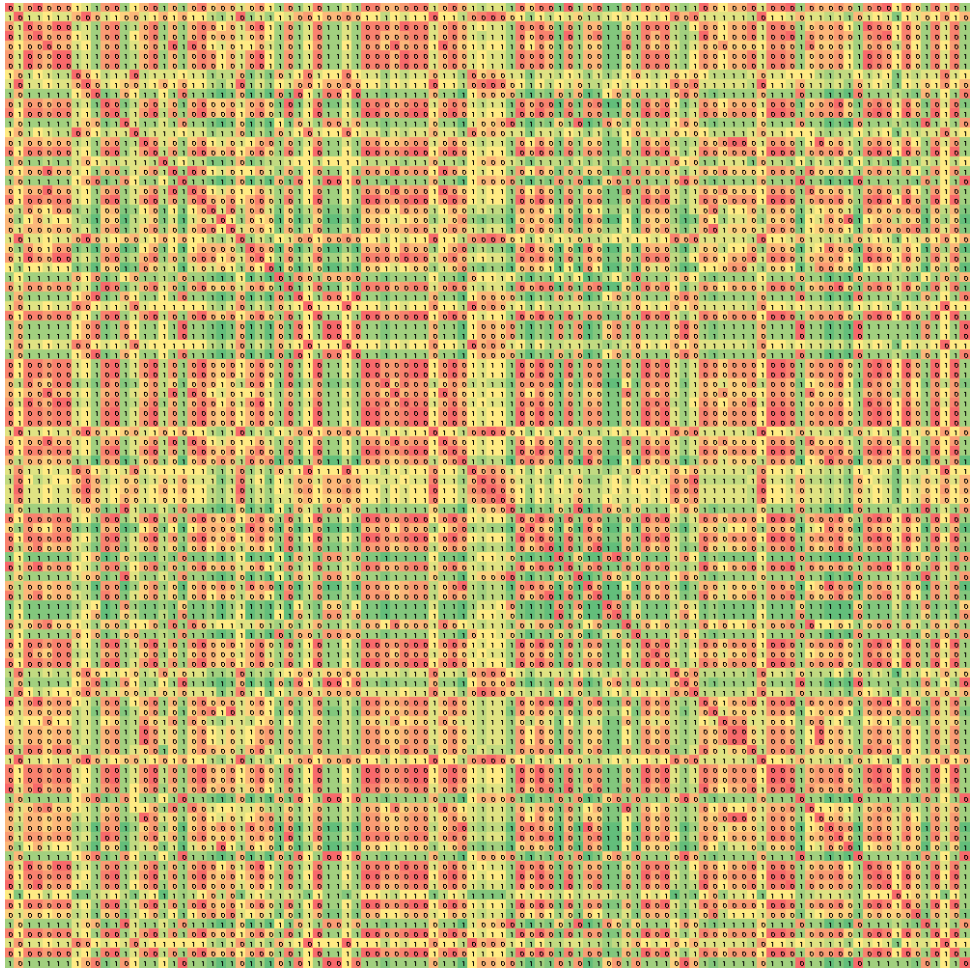
Silhouette vs K para Top20 random subsets

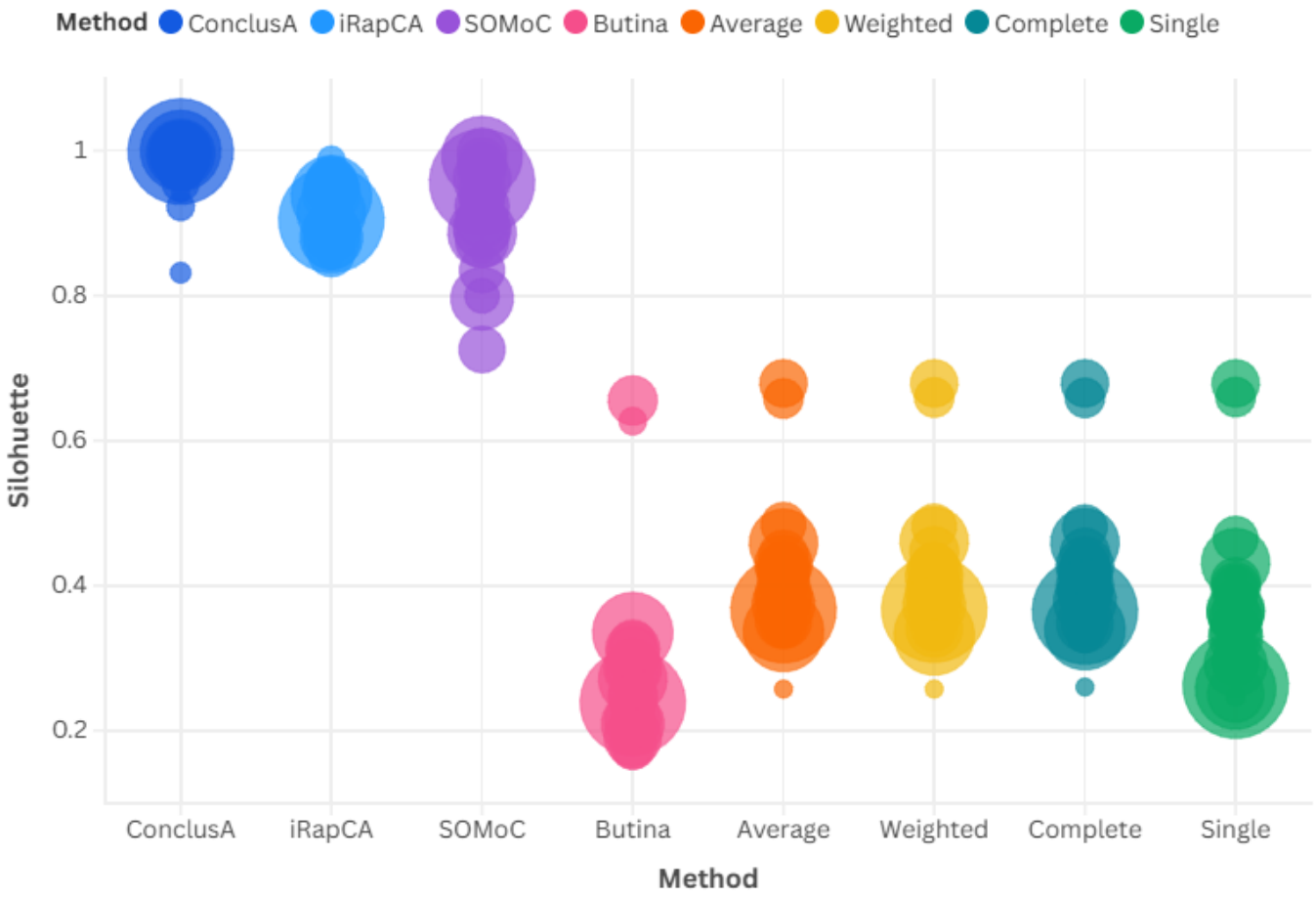


ConClusA

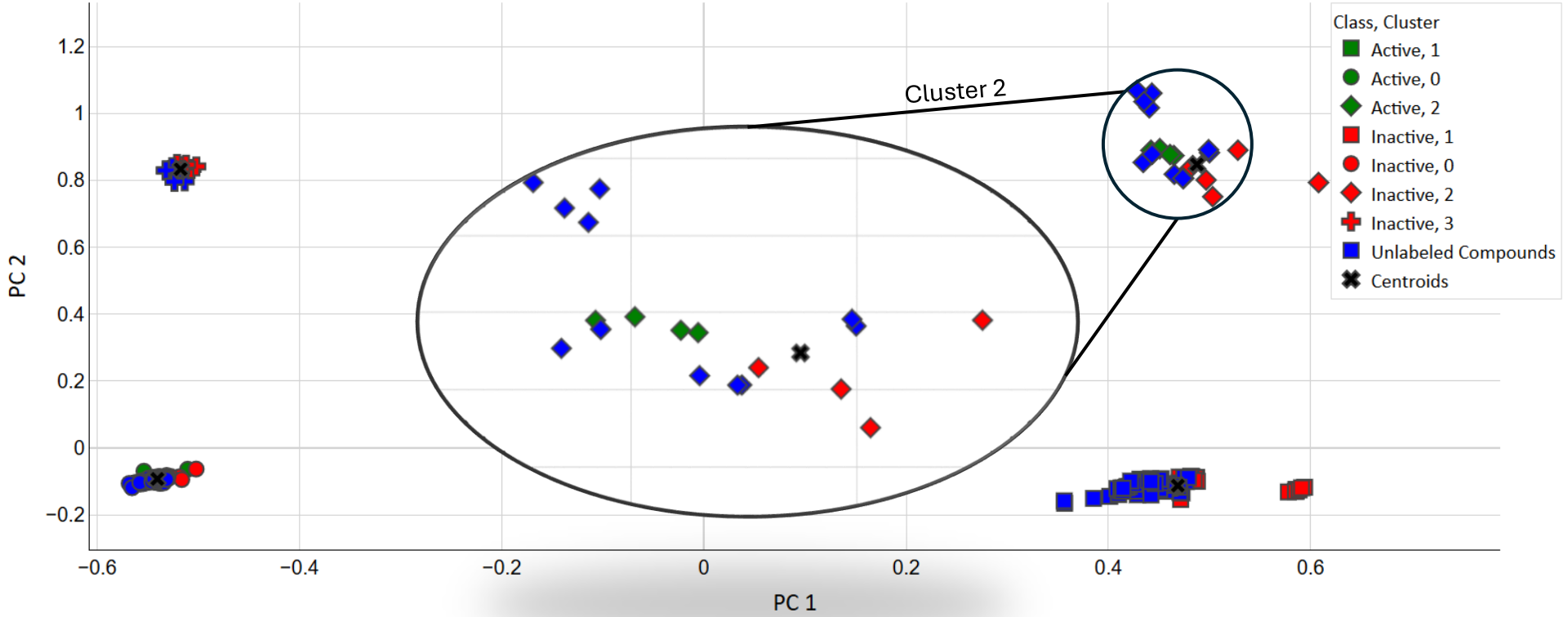







Method
ConclusA
iRapCA
SOMoC
Butina
Average
Weighted
Complete
Single
Silohuette
1
0.8
0.6
0.4
0.2
ConclusA
iRapCA
SOMoC
Butina
Average
Weighted
Complete
Single
Method

Subset 492



Color key: Known active compound | Known inactive compound | X compound (e.g., a in silico hit)

$$P(A) = \frac{Se \cdot Ya}{Se \cdot Ya + (1 - Sp) \cdot (1 - Ya)}$$

P(A): positive predictive value – probability that an in silico hit will confirm the predicted activity in vitro

Se: sensitivity, true positive rate

Sp: specificity, true negative rate

Ya: yield or proportion of active compounds in the screened chemical library (Ya is unknown *a priori*)

What should I compare the performance of a virtual screening protocol to?

In HTS campaigns, the empirical Ya (for non-focuse chemical libraries) averages 0.002 – 0.004

P(A) = 0.23 (Sp = 0.99 y Se = 1)

Samrat et al. Emerg Microbes Infect. 2023, 12, 2204164
Wu et al. Bioorg Chem. 2023, 139, 106726
Rath P et al. Nat Commun. 2023, 13, 5648
Aseeri et al. J Enzyme Inhib Med Chem 2023, 38, 343-348
Perveen et al. SLAS Discov. 2021, 26, 620-627

Antimalarial candidates

The Problem
Our Solution
Our Services
Timeline
Partners
Our team
Founders
Recognitions
Publications
Proposal

frontiers in Chemistry

ORIGINAL RESEARCH
published: 06 August 2019
doi: 10.3389/fchem.2019.00534

In silico Guided Drug Repurposing: Discovery of New Competitive and Non-competitive Inhibitors of Falcipain-2

Lucas N. Alberca [1], Sara R. Chuguransky [1], Cora L. Álvarez [2], Alan Talevi [1] and Emir Salas-Sarduy [3]*

[1] Laboratory of Bioactive Compounds Research and Development (LIDeB), Department of Biological Sciences, Exact Sciences College, Universidad Nacional de La Plata, La Plata, Argentina, [2] Departamento de Biodiversidad y Biología Experimental, Facultad de Farmacia y Bioquímica, Facultad de Ciencias Exactas y Naturales, Consejo Nacional de Investigaciones Científicas y Técnicas, Instituto de Química y Físico-Química Biológicas (IQUIFIB) "Prof. Alejandro C. Paladini", Universidad de Buenos Aires, Buenos Aires, Argentina, [3] Instituto de Investigaciones Biotecnológicas "Dr. Rodolfo Ugalde", Universidad Nacional de San Martín, CONICET, Buenos Aires, Argentina

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

Tetracycline Derivatives Inhibit Plasmodial Cysteine Protease Falcipain-2 through Binding to a Distal Allosteric Site

Jorge Enrique Hernández González, Lucas N. Alberca, Yordanka Masforrol González, Osvaldo Reyes Acosta, Alan Talevi,* and Emir Salas-Sarduy*

Cite This: *J. Chem. Inf. Model.* 2022, 62, 159−175 Read Online

Instituto de Investigaciones Biotecnológicas IIB_UNSAM

unesp

VMPaM

IICIGB CENTRO DE INGENIERÍA GENÉTICA Y BIOTECNOLOGÍA

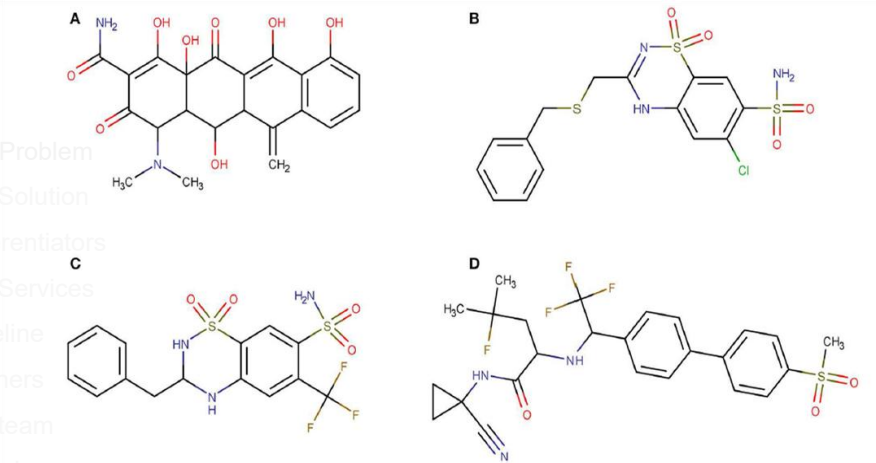


FIGURE 4 | Molecular structures of the hits selected in the prospective virtual screening campaign that were submitted to experimental confirmation. **(A)** methacycline; **(B)** benzthiazide; **(C)** bendroflumethiazide; **(D)** odanacatib.

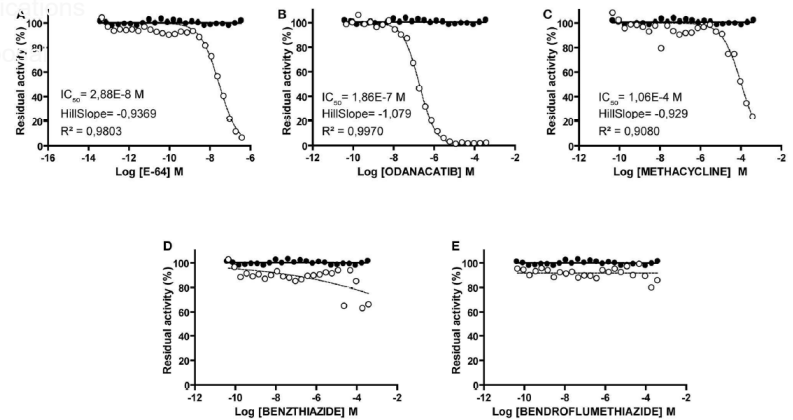


FIGURE 5 | Dose-response curves of identified falcipain-2 inhibitors (open circles). For each compound, dotted line represents the best fit of experimental data to the four-parameter Hill equation. **(A)** E-64. **(B)** Odanacatib. **(C)** Methacycline. **(D)** Benzthiazide. **(E)** Bendroflumethiazide. For those compounds achieving data convergence, the resultant values for the parameters IC50, Hillslope and R2 are indicated. In all cases, equivalent volumes of DMSO vehicle were assayed in parallel (closed circles).

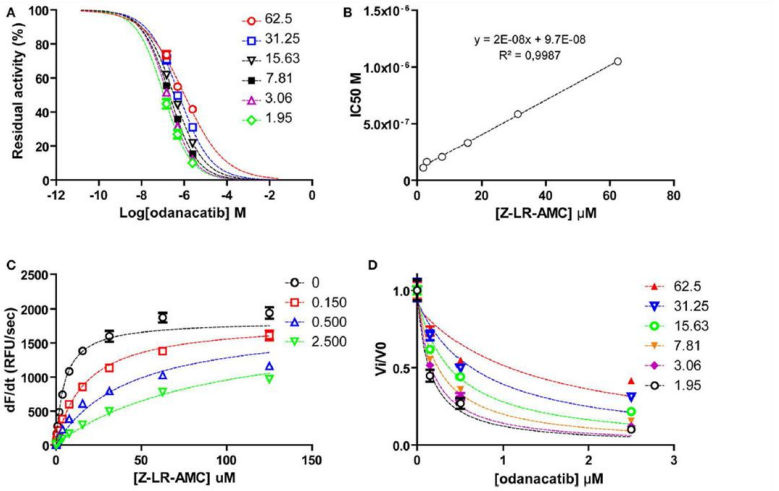


FIGURE 7 | (A) Dose-response curves for odanacatib at fixed substrate concentrations. Dotted lines represent the best fit of experimental data to the four-parameter Hill equation. **(B)** Effect of substrate concentration on the IC50 values of falcipain-2 inhibition by odanacatib. IC50 values increase linearly (~9-fold) with substrate concentration in the range 1.95–62.5 µM. Dotted line represents the best fit of data to linear equation. Y-axis intercept accounts for the Ki value. **(C)** Global fitting of kinetic data to the competitive inhibition model equation. **(D)** Global fit ting of kinetic data to the Morrison equation.

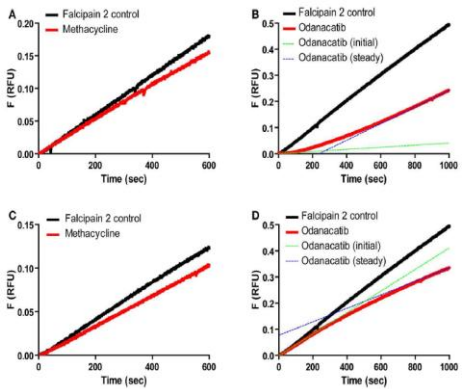


FIGURE 6 | Reversibility and time dependence of the inhibition of falcipain-2 by methacycline and odanacatib. Top panel: Product progress curves for the dissociation of E-I complex by rapid dilution (100-fold) of enzyme-inhibitor mix into substrate solution. **(A)** Methacycline. **(B)** Odanacatib. Bottom Panel: Product progress curves for the formation of E-I complex by rapid addition of enzyme to a substrate-inhibitor mix. **(C)** Methacycline. **(D)** Odanacatib.

Multi-target antiseizure drugs

The Problem
Our Solution
Differentiators
Our Services
Timeline
Partners
Our team
Founders
Recognitions
Publications
Proposal

VNiVERSiDAD D SALAMANCA

IIFP La Plata UNLP | CONICET

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

A Combined Ligand- and Structure-Based Virtual Screening To Identify Novel NaV1.2 Blockers: In Vitro Patch Clamp Validation and In Vivo Anticonvulsant Activity

Manuel A. Llanos, Nicolas Enrique, Vega Esteban-López, Sebastian Scioli-Montoto, David Sánchez-Benito, Maria E. Ruiz, Veronica Milesi, Dolores E. López, Alan Talevi, Pedro Martin, and Luciana Gavernet*

Cite This: https://doi.org/10.1021/acs.jcim.3c00645 Read Online

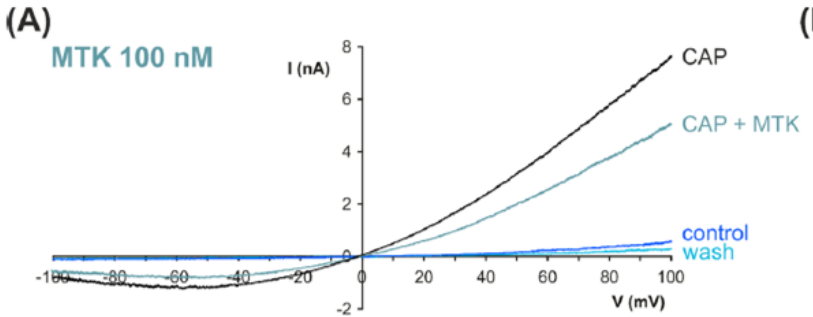
JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

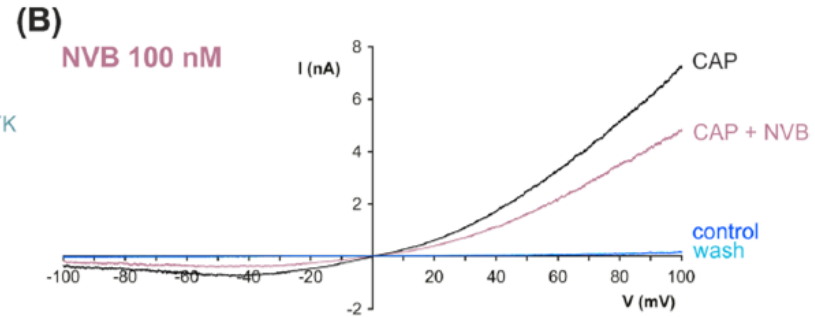
pubs.acs.org/jcim Article

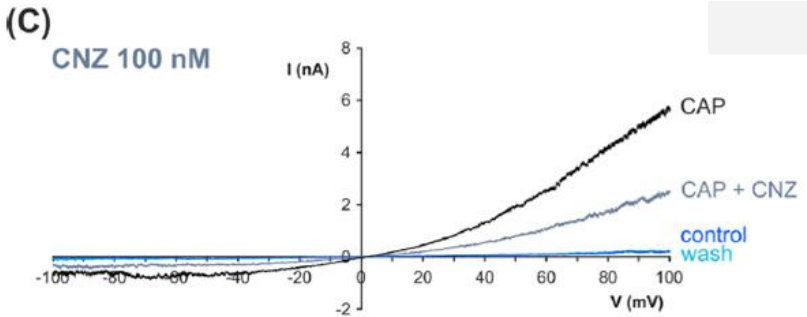
Structure-Based Virtual Screening Identifies Novobiocin, Montelukast, and Cinnarizine as TRPV1 Modulators with Anticonvulsant Activity *In Vivo*

Manuel A. Llanos,○ Nicolás Enrique,○ María L. Sbaraglini, Federico M. Garofalo, Alan Talevi, Luciana Gavernet,* and Pedro Martín

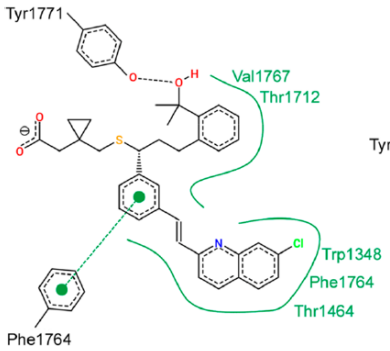
Cite This: https://doi.org/10.1021/acs.jcim.2c00312 Read Online





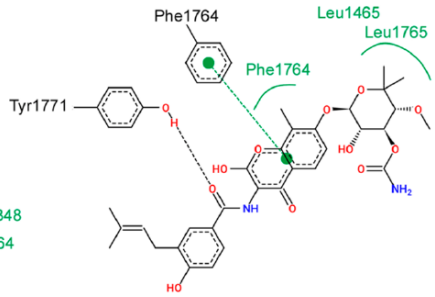


Montelukast



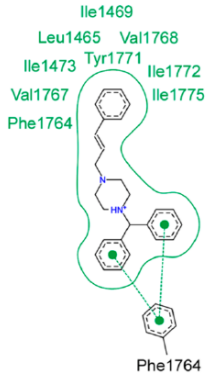
QSAR: 5.75 | **Docking:** -11.63

Novobiocin

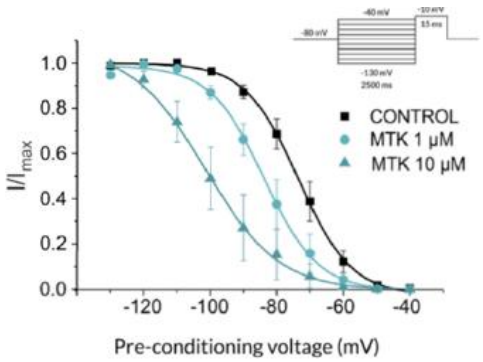


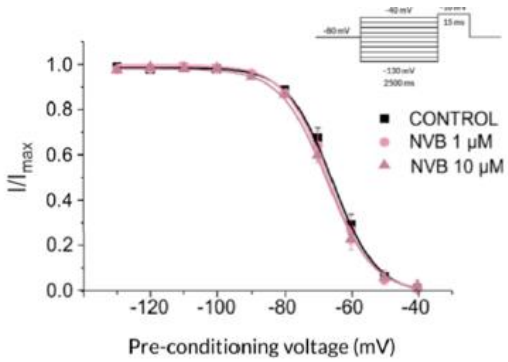
QSAR: 5.38 | **Docking:** -12.68

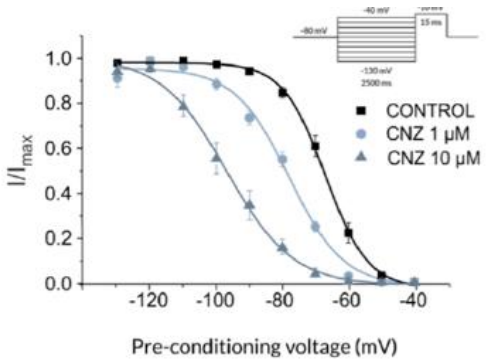
Cinnarizine



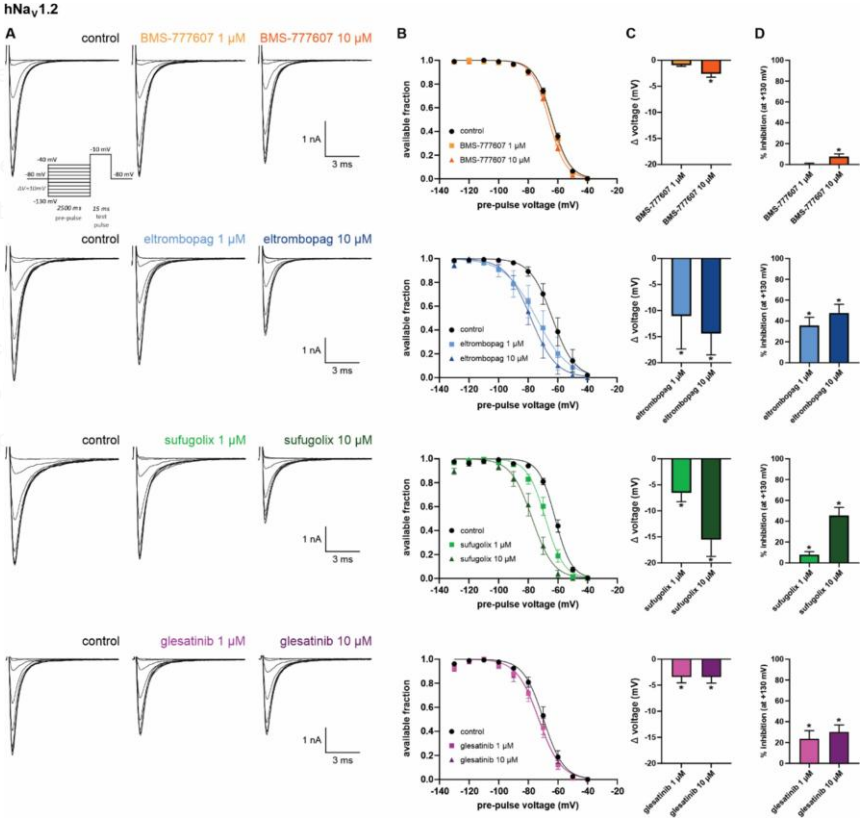
QSAR: 6.44 | **Docking:** -9.12







Sodium channel inhibitors



Brain Research 1856 (2025) 149571

Contents lists available at ScienceDirect

Brain Research

journal homepage: www.elsevier.com/locate/brainres

Research paper

In silico screening to search for selective sodium channel blockers: When size matters☆

Maximiliano José Fallico [a], Lucas Nicolás Alberca [a], Nicolás Enrique [b], Federico Orsi [b], Denis Nihuel Prada Gori [a], Pedro Martín [b], Luciana Gavernet [a], Alan Talevi [a,*]

Drugs against Chagas disease

JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

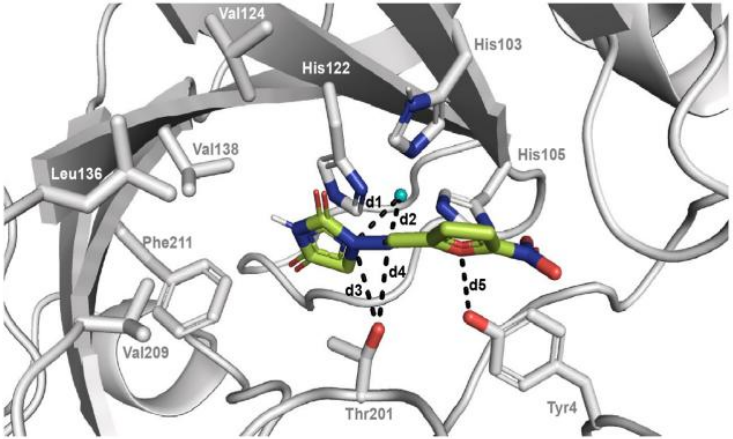
pubs.acs.org/jcim Article

Discovery of *Trypanosoma cruzi* Carbonic Anhydrase Inhibitors by a Combination of Ligand- and Structure-Based Virtual Screening

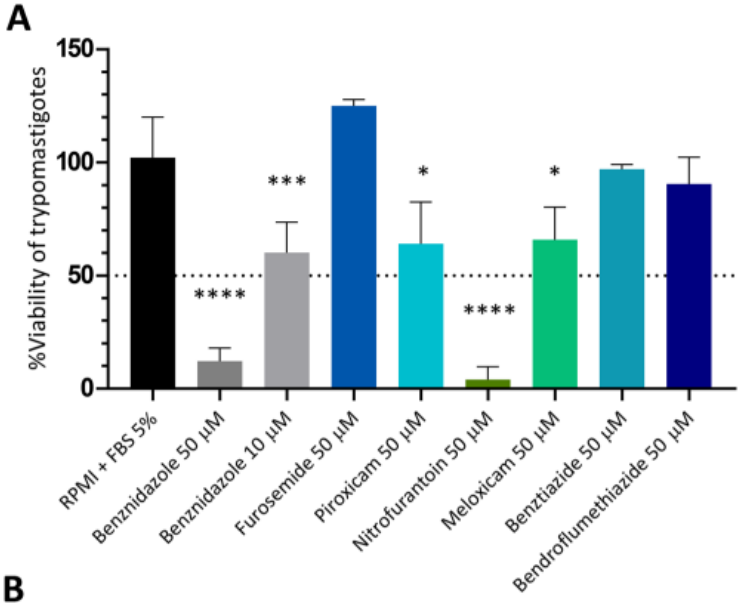
Denis N. Prada Gori, Emilia M. Barrionuevo, Lucas N. Alberca, María L. Sbaraglini, Manuel A. Llanos, Simone Giovannuzzi, Fabrizio Carta, Matías I. Marchetto, Claudiu T. Supuran, Catalina D. Alba Soto, Luciana Gavernet, and Alan Talevi*

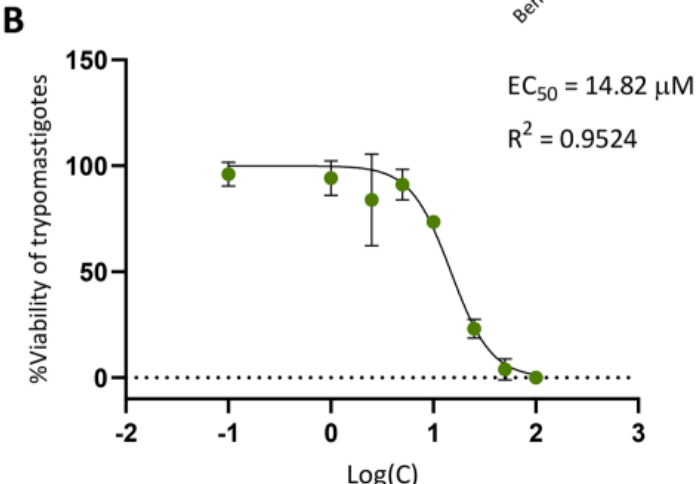
iMPaM

FLORENTINA STUDIORUM UNIVERSITAS



name	K_i (nM)[b]		selectivity ratio
	*Tc*CA	hCA II	
benzthiazide	934	9[c]	9.4×10^{-3}
bendroflumethiazide	728	1700[c]	2.3
furosemide	394	65[c]	0.2
meloxicam	3831	6430[c]	1.7
piroxicam	6001	6100[c]	1.0
nitrofurantoin	93	73	0.8
acetazolamide	62	12	0.2





SARS-CoV2 mPro inhibitors

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 22 June 2023
DOI 10.3389/fphar.2023.1193282

OPEN ACCESS

EDITED BY
Amit Khurana,
University Hospital RWTH Aachen, Germany

REVIEWED BY
Priyanka Banerjee,
Charité University Medicine Berlin, Germany
Kittikhun Wangkanont,
Chulalongkorn University, Thailand
Paola Storici,
Elettra Sincrotrone Trieste, Italy

*CORRESPONDENCE
Alan Talevi,
alantalevi@gmail.com
Marcelo A. Comini,
mcomini@pasteur.edu.uy

†Deceased

‡These authors have contributed equally to this work and share first authorship

¶These authors have contributed equally to this work and share senior authorship

RECEIVED 24 March 2023
ACCEPTED 31 May 2023
PUBLISHED 22 June 2023

CITATION
Ruatta SM, Prada Gori DN, Fló Díaz M, Lorenzelli F, Perelmuter K, Alberca LN, Bellera CL, Medeiros A, López GV, Ingold M, Porcal W, Dibello E, Ihnatenko I, Kunick C, Incerti M, Luzardo M, Colobbio M, Ramos JC, Manta E, Minini L, Lavaggi ML, Hernández P, Šarlauskas J, Huerta García CS, Castillo R, Hernández-Campos A, Ribaudo G, Zagotto G, Carlucci R, Medrán NS, Labadie GR, Martínez-Amezaga M,

Garbage in, garbage out: how reliable training data improved a virtual screening approach against SARS-CoV-2 MPro

Santiago M. Ruatta[1,2†], Denis N. Prada Gori[3†], Martín Fló Díaz[4,5†], Franca Lorenzelli[1], Karen Perelmuter[6], Lucas N. Alberca[3,7], Carolina L. Bellera[3,7], Andrea Medeiros[1,8], Gloria V. López[9,10], Mariana Ingold[10], Williams Porcal[9,10], Estefanía Dibello[9], Irina Ihnatenko[11], Conrad Kunick[11], Marcelo Incerti[9], Martín Luzardo[9], Maximiliano Colobbio[9,12], Juan Carlos Ramos[9,12], Eduardo Manta[9,12], Lucía Minini[9], María Laura Lavaggi[13], Paola Hernández[14], Jonas Šarlauskas[15], César Sebastian Huerta García[16], Rafael Castillo[16], Alicia Hernández-Campos[16], Giovanni Ribaudo[17], Giuseppe Zagotto[18], Renzo Carlucci[19], Noelia S. Medrán[19], Guillermo R. Labadie[19], Maitena Martinez-Amezaga[19], Carina M. L. Delpiccolo[19], Ernesto G. Mata[19], Laura Scarone[9], Laura Posada[9], Gloria Serra[9], Theodora Calogeropoulou[20], Kyriakos Prousis[20], Anastasia Detsi[21], Mauricio Cabrera[22], Guzmán Alvarez[22], Adrián Aicardo[8,23,24], Verena Araújo[8,23,25], Cecilia Chavarría[8,23], Lucija Peterlin Mašič[26], Melisa E. Gantner[3,7], Manuel A. Llanos[3,7], Santiago Rodríguez[3], Luciana Gavernet[3,7], Soonju Park[27], Jinyeong Heo[27], Honggun Lee[27], Kyu-Ho Paul Park[27], Mariela Bollati-Fogolín[6], Otto Pritsch[4,5§], David Shum[27], Alan Talevi[3,7*‡] and Marcelo A. Comini[1*‡]

PASTEUR NETWORK

Institut Pasteur de Montevideo

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

MERCOSUR FOCEM
Fondo para la Convergencia Estructural del Mercosur

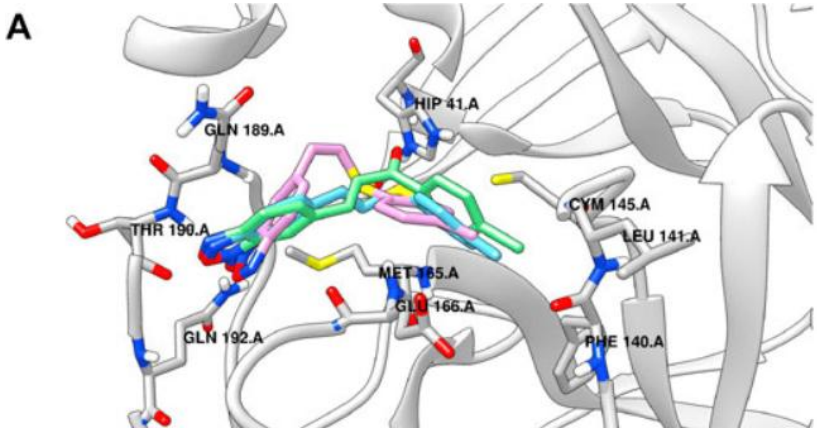
CONICET

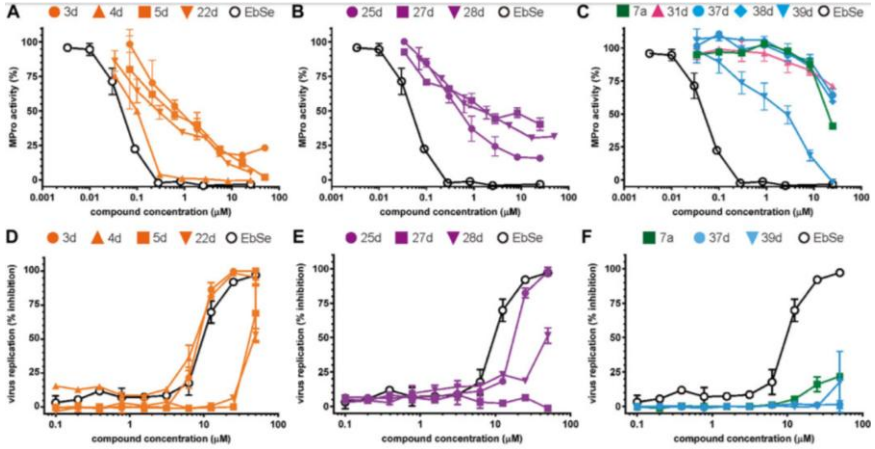
NRF 한국연구재단 National Research Foundation of Korea

DFG Deutsche Forschungsgemeinschaft
German Research Foundation

ICGEB

CONACYT





Total screened compounds

6,266 compounds

43

Predicted in silico hits within applicability domain

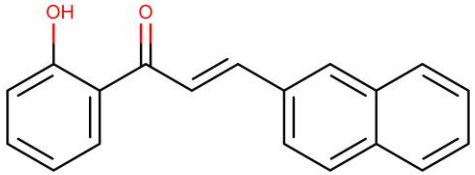
Available compounds submitted to confirmation

28

8

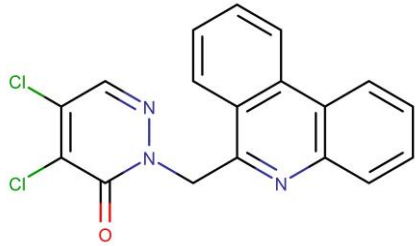
Confirmed hits (28,6%) with three different scaffolds

IC_{50} in the range 0.46-5.76 μM

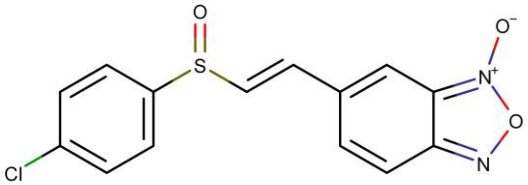


IC_{50} = 0.46 +/- 0.21

EC_{50} = 17.5



IC_{50} = 1.97 +/- 0.50



IC_{50} = 0.121 +/- 0.002

EC_{50} = 7.3

Thanks!

Questions?

EMAIL ADDRESS
alantalevi@gmail.com
atalevi@biol.unlp.edu.ar

WEBSITES
www.lideb.biol.unlp.edu.ar
www.boolzi.com

LIDeB
Laboratorio de Investigación y Desarrollo de Bioactivos

Boolzi
Forging tomorrow's cures