

L'uso del linguaggio di programmazione Python per la costruzione di un database di molecole di origine naturale

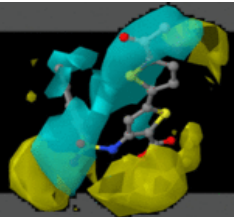


SAPIENZA
UNIVERSITÀ DI ROMA

**Facoltà di Farmacia e Medicina
Corso di Laurea in Chimica e Tecnologia Farmaceutiche
Tesi Sperimentale in Chimica Farmaceutica
a.a. 2012/2013**

**Laureando: Eleonora Menicacci
Matricola: 1332005**

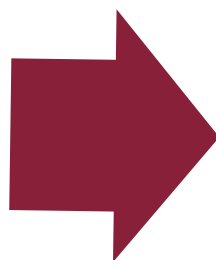
Relatore: prof. Rino Ragno



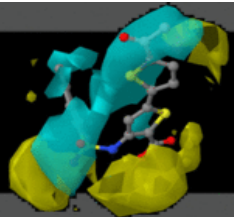
INTRODUZIONE

by www.RCMD.it

SCOPO



COSTRUZIONE
DI UN
DATABASE DI
MOLECOLE DI
ORIGINE
NATURALE



perché la necessità di
questo database?

?

?

?

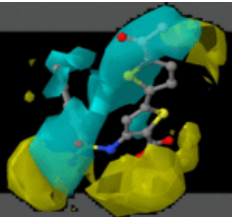
?

?

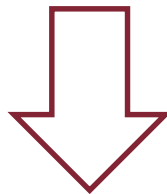
?

?

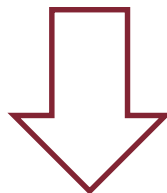
?



Crescente necessità di catalogare dati



Diffusione di banche dati chimiche



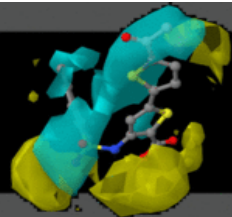
Integrazione dell'
informatica nelle scienze
chimiche e biologiche



**Viene descritta
l'utilità del linguaggio
di programmazione
Python nella
biologia, astrofisica e
neuroscienze**



**Viene finalmente dato un
riconoscimento ufficiale alla
CHEMIOINFORMATICA**



The Nobel Prize in Chemistry 2013

Martin Karplus, Michael Levitt, Arieh Warshel

The Nobel Prize in Chemistry 2013



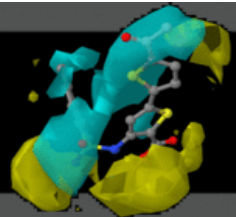
© Nobel Media AB
Martin Karplus



Photo: Keilana via
Wikimedia Commons
Michael Levitt



Photo: Wikimedia
Commons
Arieh Warshel

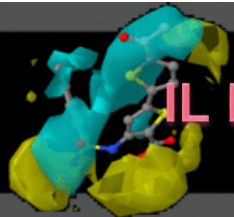


2013: premio Nobel per la chimica dato a tre
Chimici Computazionali:

Michael Levitt, Martin Karplus, Arieh Warshel
che hanno condotto studi sulla molecola del

RETINALE

sviluppando un programma in grado di includere
le leggi della meccanica quantistica
per lo studio del comportamento
degli elettroni



Python



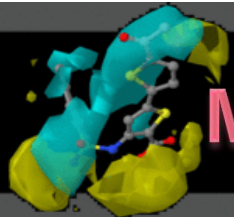
linguaggio potente

semplice da studiare

facile utilizzazione

elaborare dati in modo efficiente e veloce

opensource



Strumenti
impiegati

Python

Django

Openmoldb

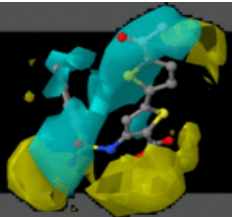
Openbabel e pybel

Rdkit

Jsme e Jsmol

Chemdoodle

PubChemPy



RISULTATI

by www.RCMD.it

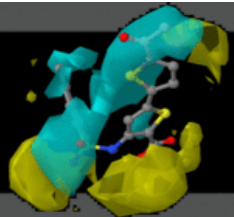
Il risultato ottenuto è un web server pienamente funzionante

L'obiettivo futuro del laboratorio RCMD è implementare questo server con nuove applicazioni:

Allineamenti molecolari

Calcoli 3-D_QSAR

Applicazioni structure-based (SB, molecular docking)



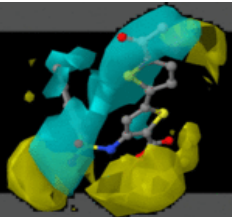
RISULTATI

by www.RCMD.it

Per il momento è disponibile solo il database contenente molecole di origine naturale.

La compilazione del database ha richiesto un lavoro pionieristico e inconsueto per un corso di laurea imperniato nelle scienze farmaceutiche.

Tuttavia i recenti report scientifici su riviste del settore farmaceutico riportano sempre più metodi di “drug design” direttamente implementati in web server.

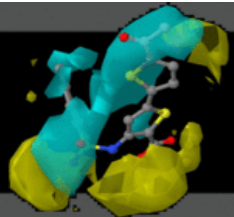


VANTAGGI

by www.RCMD.it

Ricerca attività farmacologica di
composti naturali

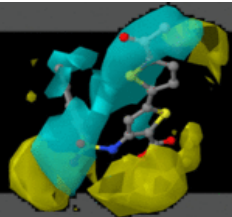
Possibilità di catalogazione
sistematica di composti



Il Server

by www.RCMD.it

Come si presenta all'interno il server?



II Server

by **www.RCMD.it**

Home page

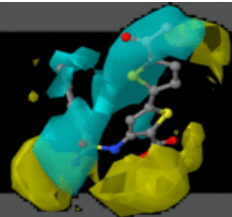


Rcmd's Natural Databank

Substance libraries, composed of million of compounds, are tested in the pharmaceutical industry with robotic systems. Furthermore, the follow-up of the "hits" is often so expensive that essentially only large companies can use this method. Virtual screening helps to decide which compounds to screen, which libraries to synthesis and which compounds to purchase from an external supplier reducing the overall cost associated to the discovery and development of new drugs.

Main classes of virtual screening methods are:

- Similarity search (ligand-based virtual

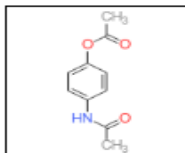


II Server

by **www.RCMD.it**

BrowseDB

RCMD1

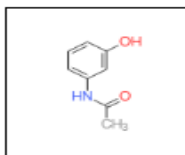


Molecular Weight	LogP	Carbon Atoms	Rotable Bonds	TPSA	Num HDonors	Num HAcceptors
193.202	1.5703	14	4	55.4	1	3

SMILES: CC(=O)Nc1ccc(cc1)OC(=O)C

[Download chemical-physical calculated properties in a csv file](#)

RCMD2

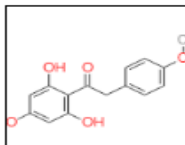


Molecular Weight	LogP	Carbon Atoms	Rotable Bonds	TPSA	Num HDonors	Num HAcceptors
151.165	1.3506	11	2	49.33	2	2

SMILES: CC(=O)Nc1cccc(c1)O

[Download chemical-physical calculated properties in a csv file](#)

RCMD3



Molecular Weight	LogP	Carbon Atoms	Rotable Bonds	TPSA	Num HDonors	Num HAcceptors
274.272	2.2374	20	4	86.99	3	5

SMILES: COc1ccc(cc1)CC(=O)c2c(cc(cc2O)O)O

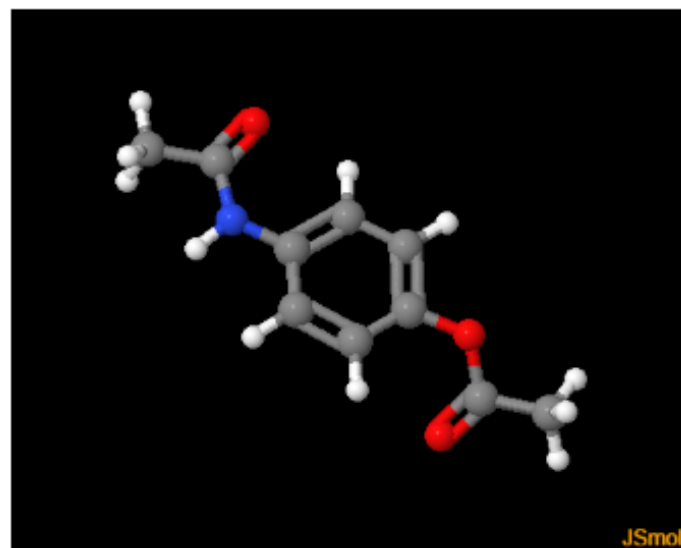
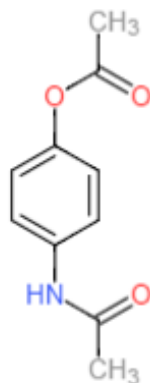


II Server

by www.RCMD.it

Molecule

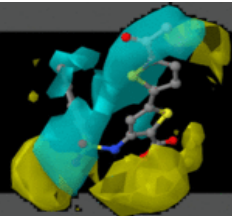
 You are inspecting RCMD1



JSmol

Molecular Weight	LogP	Carbon Atoms	Rotable Bonds	TPSA	Num HDonors	Num HAcceptors	CID
193.202	1.5703	14	4	55.4	1	3	17499

Click on Cid number to connect
PubChem for the RCMD1
molecule
Cid: [17499](#)



II Server

by www.RCMD.it

Search

For the search you need to export chemical structure from JSME applet:

[Click Here to Get SMILES Code](#)

Filtering

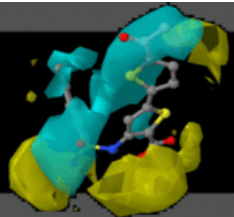
Molecular Weight:
LogP:

TPSA:
Hydrogen Acceptors:
Hydrogen Donors:
Tanimoto:

default Lipinski's rule of five ranges

[Click Here to Get Tanimoto Similarity Based Compounds](#)

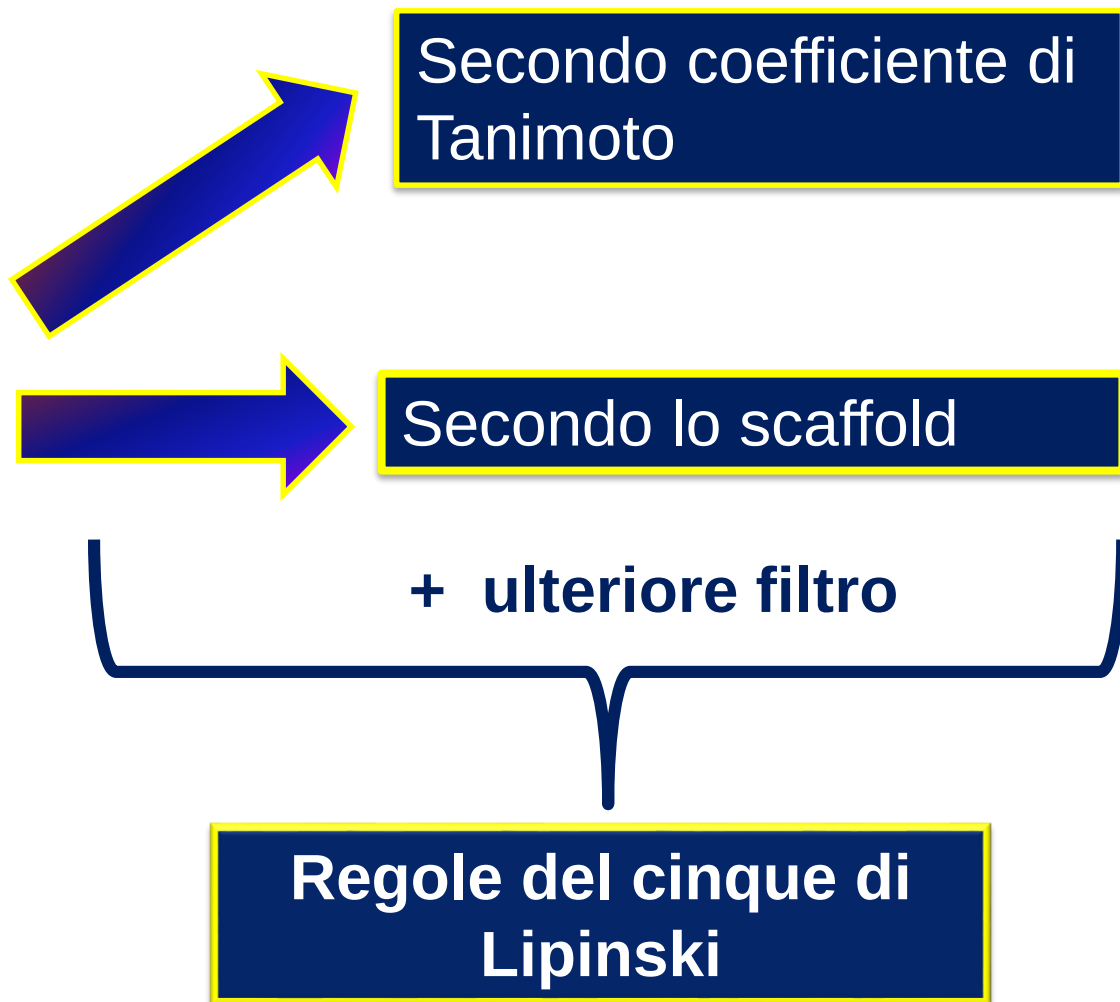
[Click Here to Get Substructure Containing Compounds](#)

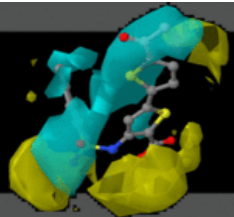


FILTRI DELLA FUNZIONE SEARCH

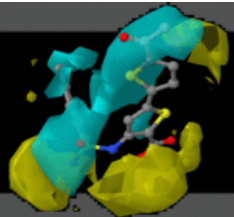
by www.RCMD.it

Questo database offre la possibilità di ricercare le molecole





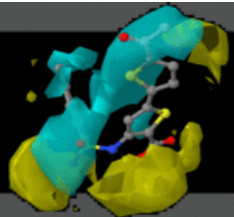
È un numero compreso tra 0 e 1
Calcola la similarità tra due o
più molecole



scaffold

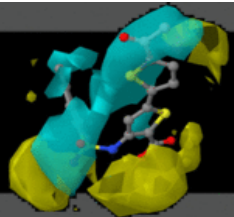
by www.RCMD.it

Si vanno a ricercare tutte le molecole contenenti al loro interno la struttura disegnata



Una
molecola, per
essere
“druggable”
DEVE

- Rispettare
le regole
di Lipinski



tutti multipli di 5!

4 REGOLE

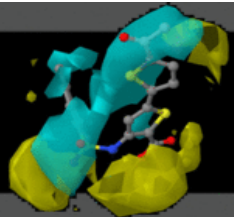
Ulteriore filtro: regole del 5 di Lipinski

Non più di cinque donatori di legami idrogeno

Non più di dieci accettori di legami idrogeno

Peso molecolare minore di 500 daltons

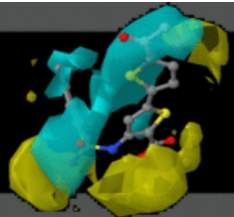
Coefficiente di partizione ottanolo-acqua non più di 5



Lipinski

by www.RCMD.it

Impostare come filtro di default i valori di Lipinski vuol dire, con una buona probabilità, ottenere molecole che abbiano un'attività biologica e/o farmacologica

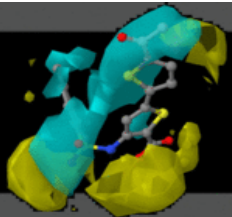


Lipinski

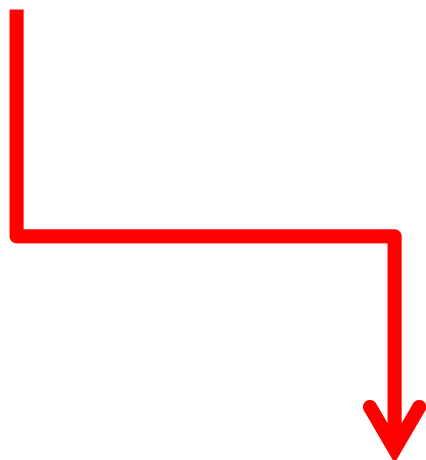
by www.rcmd.it

Impostare come filtro di default i valori di Lipinski vuol dire, con una buona probabilità, ottenere molecole che abbiano un'attività biologica e/o farmacologica

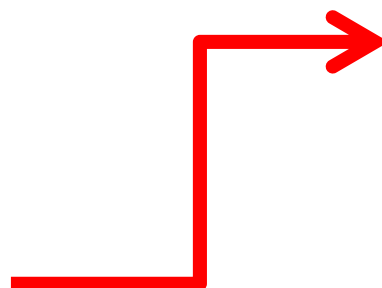
Pfizer



Come possiamo
conoscere l'attività
biologica?



CID



PUBCHEM



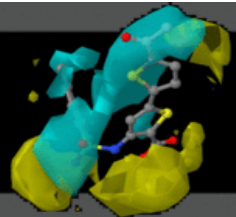
Contiene solo molecole di origine naturale, escludendo quelle di sintesi

Parte delle molecole presentano attività farmacologica, confermabile tramite collegamento diretto a PubChem

È accessibile a tutti (<http://naturaldb.rcmd.it>)

Si può individuare velocemente un set di molecole da cui si può, per similitudine con altri farmaci, identificare una potenziale attività e risalire alla struttura in 2D, 3D e a tutti i parametri

Strumento di screening veloce, utile per la scelta di molecole di cui intraprendere la sintesi, partendo da strutture presenti in natura, che possono avere un numero minore di effetti collaterali rispetto ai farmaci totalmente sintetici.



Live

by *www.RCMD.it*

<http://naturaldb.rcmd.it/>