

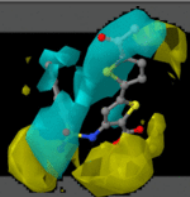
L'allineamento nel molecular modelling



SAPIENZA
UNIVERSITÀ DI ROMA

Manuela Sabatino

Lunedì
13 Novembre 2017



(3-D QSAR)

Quantitative **S**tructure-**A**ctivity **R**elationships

- Spiegazione dei dati Biologici su base tridimensionale
- Ottimizzazione molecole esistenti
- Predizione composti non ancora saggiati e/o sintetizzati



COSTRUZIONE DI UN MODELLO

Scelta del training set

Allineamento delle molecole

Calcolo dei campi di interazione molecolare (MIF)

Generazione del modello statistico

Validazione interna, esterna e y-scrambling



Scelta del training set

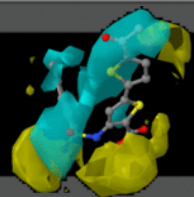
Allineamento delle molecole

Calcolo del campi di interazione molecolare (MIF)

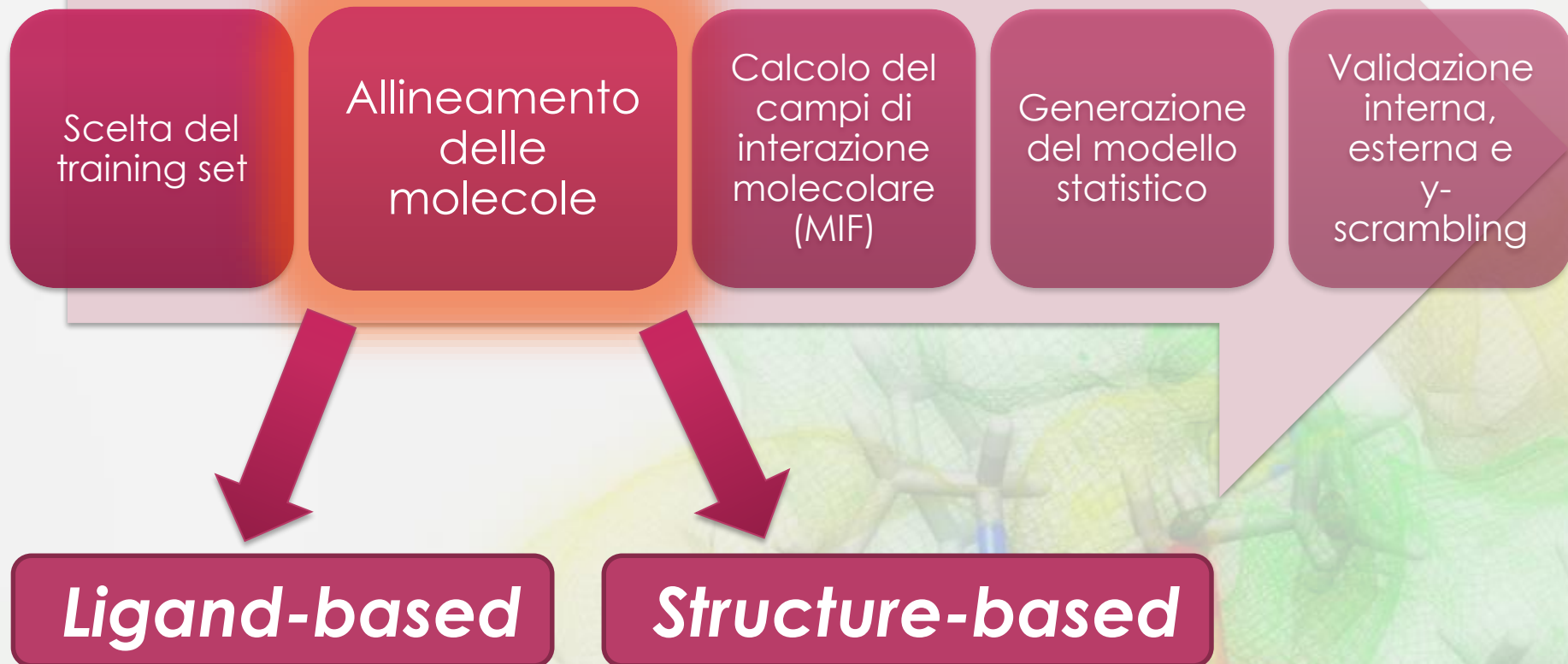
Generazione del modello statistico

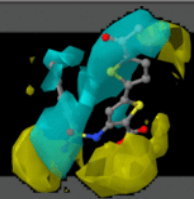
Validazione interna, esterna e y-scrambling

- Tipologie di allineamento;
- Alcuni algoritmi usati;
- Valutazione di un allineamento;
- Software che sviluppano allineamenti;



TIPOLOGIE DI ALLINEAMENTO





APPROCCIO STRUCTURE-BASED

“This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.”

Outlook.com - manu x RCSB Protein Data E x

www.rcsb.org/pdb/home/home.do

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB Login

RCSB PDB An Information Portal to 113494 Biological Macromolecular Structures

Search by PDB ID, author, macromolecule, sequence, or ligands Go

Advanced Search | Browse by Annotations

PDB-101 PDB EMDatabank Structural Biology Knowledgebase

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

A Structural View of Biology

This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

Take an Interactive Tour of the PDB

BEGIN EXPLORING

November Molecule of the Month

glutamate

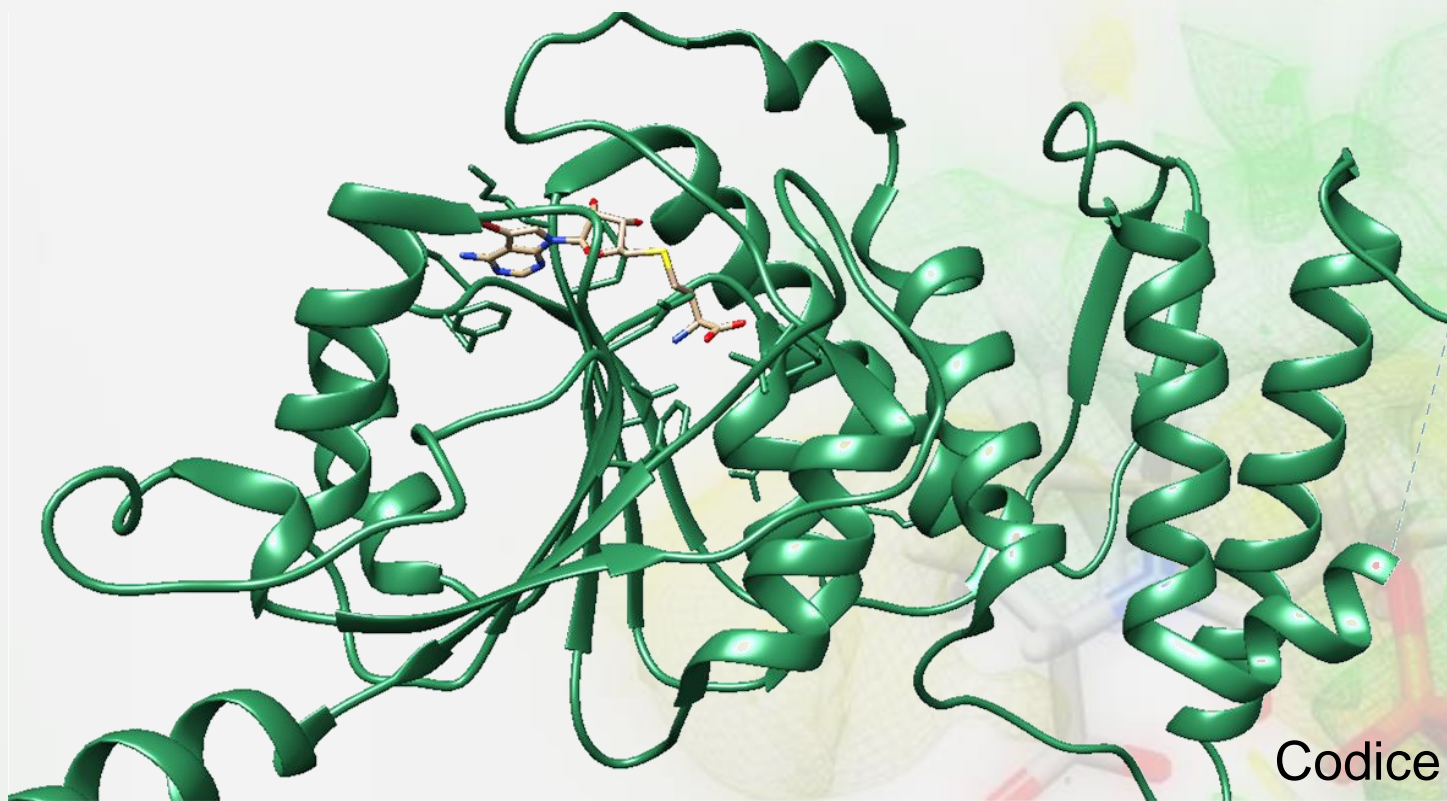
ivermectin

Glutamate-gated Chloride Receptors

Feedback



DOT1L Disruptor of Telomeric Silencing 1-like



Codice complesso 3SX0



Codice ID	Risoluzione (Å)
-----------	-----------------

1NW3	2.50
------	------

3QOW	2.10
------	------

3QOX	2.30
------	------

3SR4	2.50
------	------

3SX0	2.28
------	------

3UWP	2.05
------	------

4EK9	2.50
------	------

4EKG	2.80
------	------

4EKI	2.85
------	------

4ER5	2.57
------	------

4EQZ	2.15
------	------

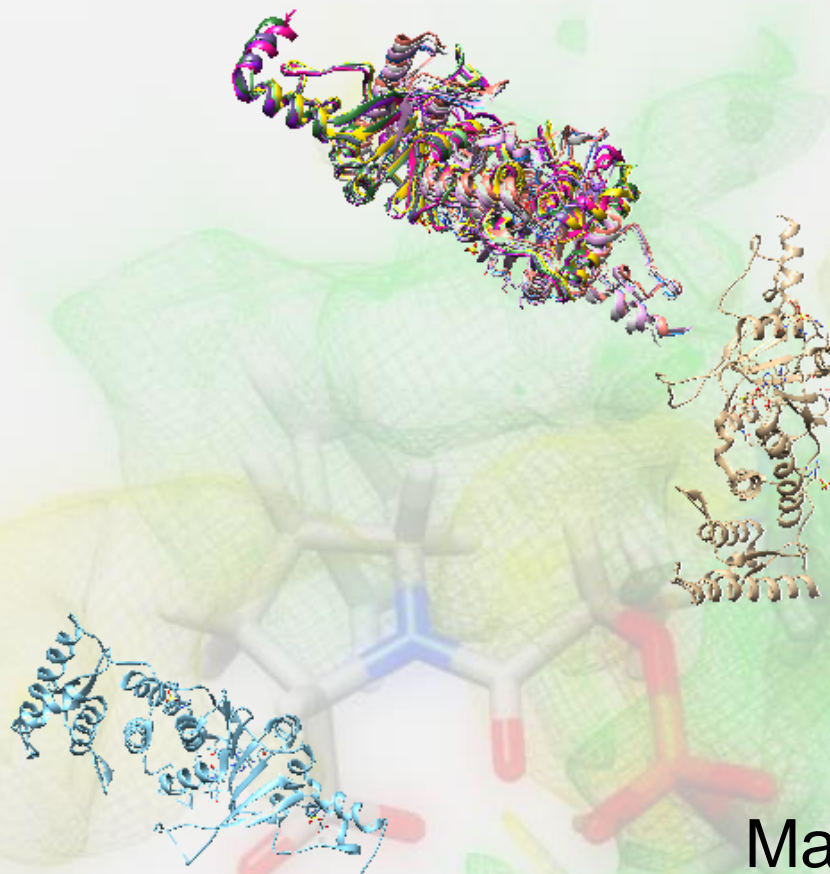
4ER0	2.50
------	------

4ER6	2.30
------	------

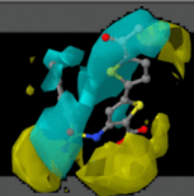
4ER7	2.20
------	------

4HRA	3.15
------	------

DOT1L Disruptor of Telomeric Silencing 1-like



MatchMaker



Codice ID	Risoluzione (Å)
-----------	-----------------

1NW3	2.50
------	------

3QOW	2.10
------	------

3QOX	2.30
------	------

3SR4	2.50
------	------

3SX0	2.28
------	------

3UWP	2.05
------	------

4EK9	2.50
------	------

4EKG	2.80
------	------

4EKI	2.85
------	------

4ER5	2.57
------	------

4EQZ	2.15
------	------

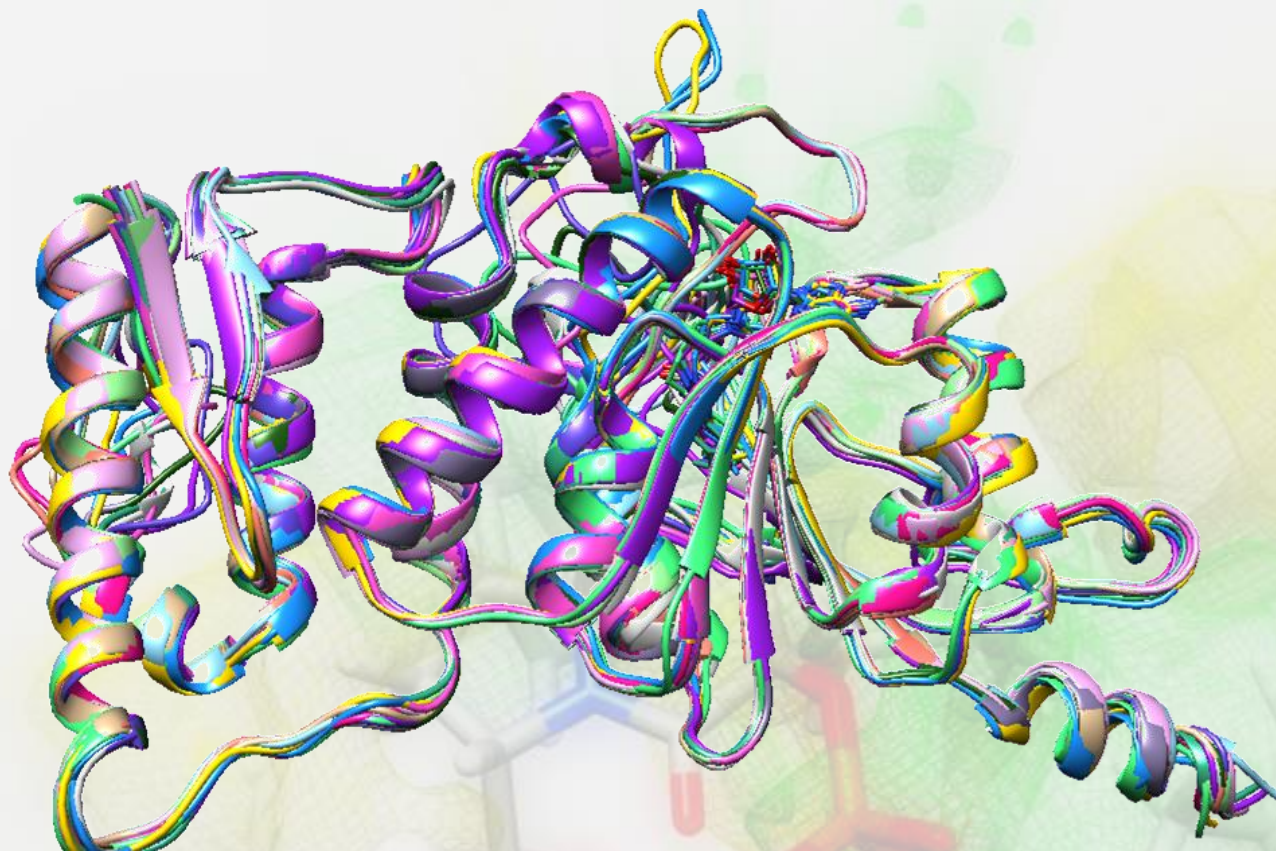
4ER0	2.50
------	------

4ER6	2.30
------	------

4ER7	2.20
------	------

4HRA	3.15
------	------

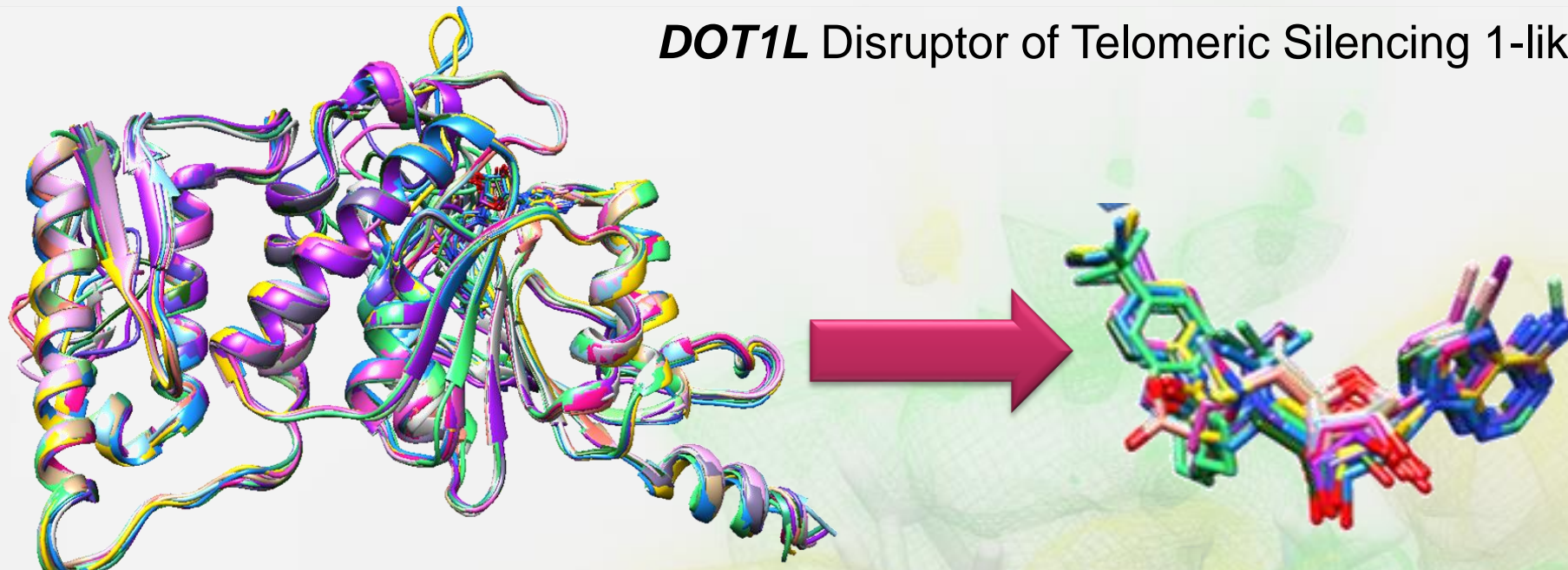
DOT1L Disruptor of Telomeric Silencing 1-like



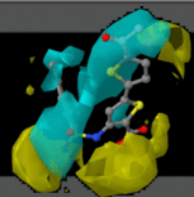
MatchMaker



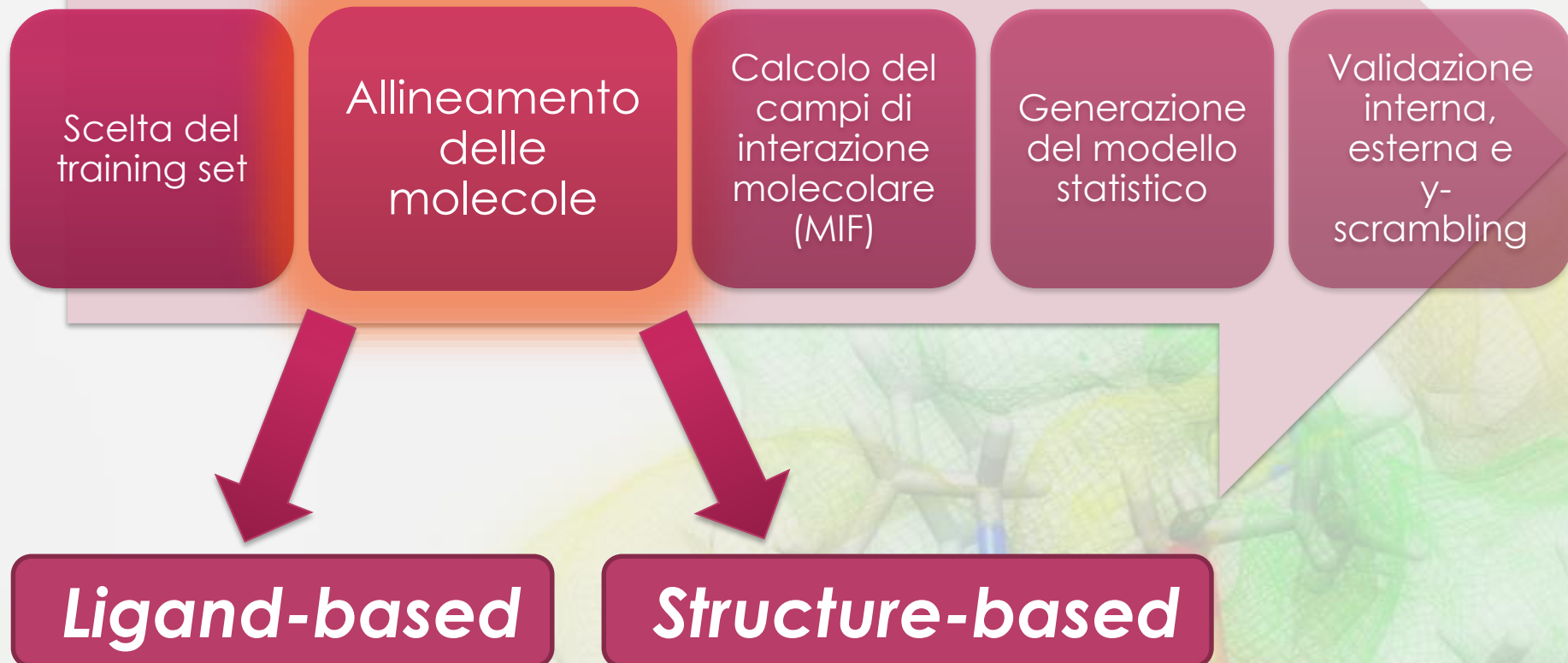
DOT1L Disruptor of Telomeric Silencing 1-like

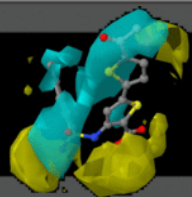


Allineamento con approccio
STRUCTURE-BASED

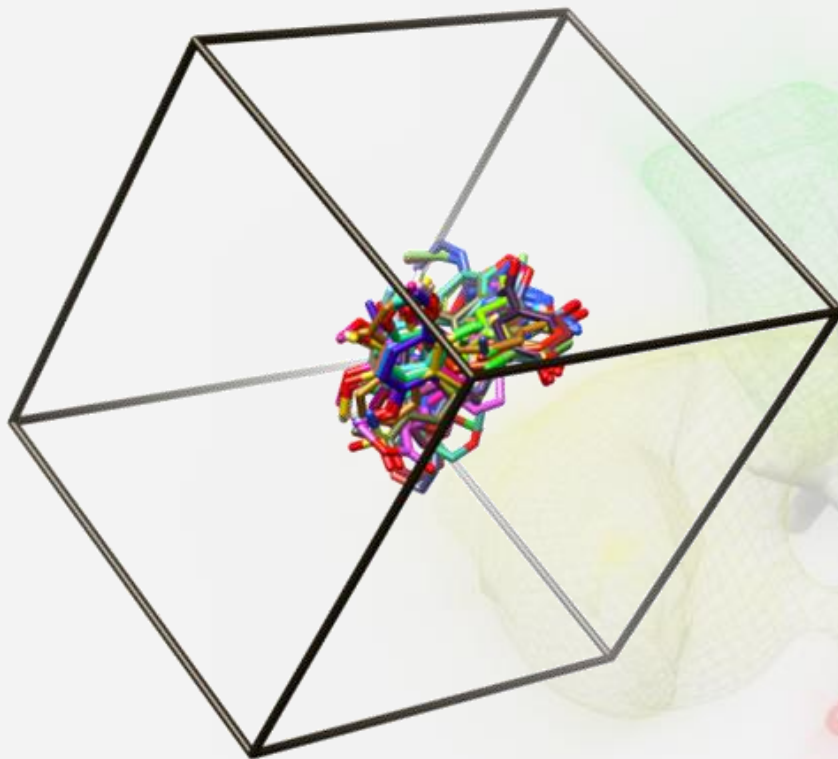


TIPOLOGIE DI ALLINEAMENTO





- ✓ Viene utilizzato in assenza di Dati 3-D del target su cui i nostri inibitori agiscono;
- ✓ Utile a individuare una disposizione spaziale del dataset che tenta di mimare una ipotetica binding mode;



- ✓ Molecola più attiva;
- ✓ Molecola meno attiva;
- ✓ Molecola più flessibile;
- ✓ Molecola meno flessibile;
- ✓ Molecola più pesante;
- ✓ Molecola più lunga e voluminosa;



I Criteri mediante il quale l'allineamento può essere sviluppato sono diversi:

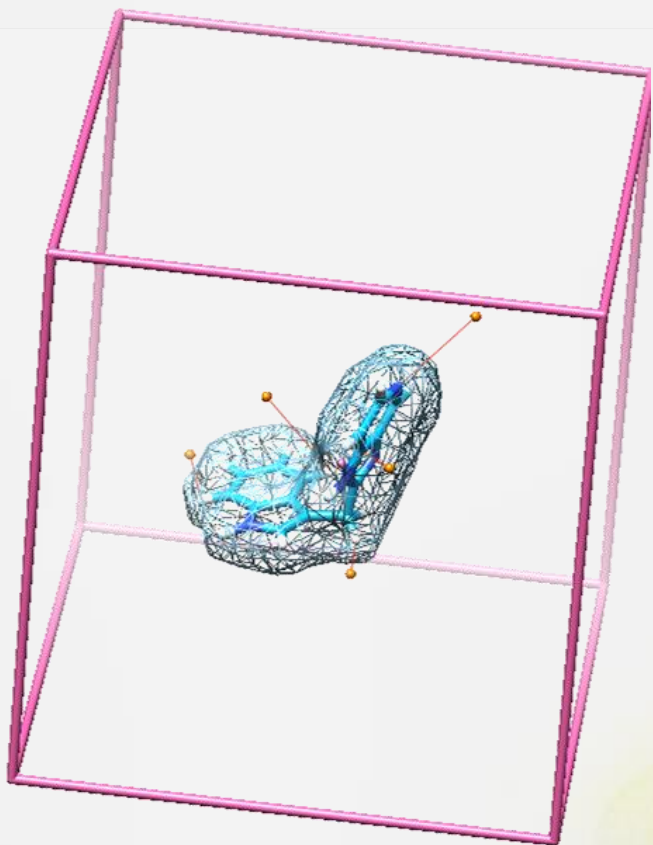
- Similarità morfologica;
 - Atomico;
 - Farmacoforico;
- campi del potenziale elettrostatico molecolare (MEP).



- Similarità morfologica (MS): prende in considerazione la differenza nelle distanze della superficie molecolare di due molecole da precisi punti osservazionali individuati in una griglia uniforme.



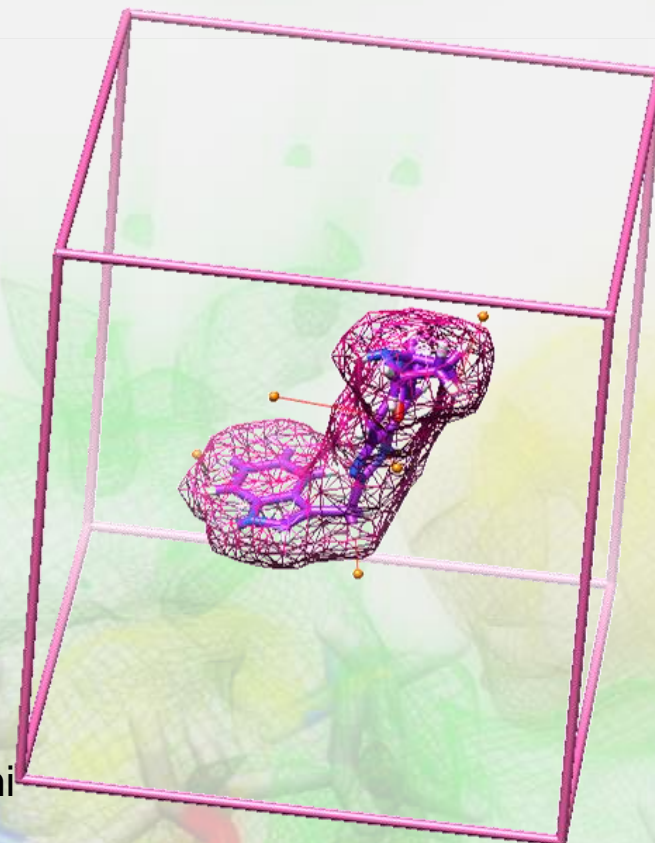
PROGRAMMI DI SOVRAPPOSIZIONE MOLECOLARE : SURFLEX



➤ Distanza del punto osservazionale dalla superficie VdW

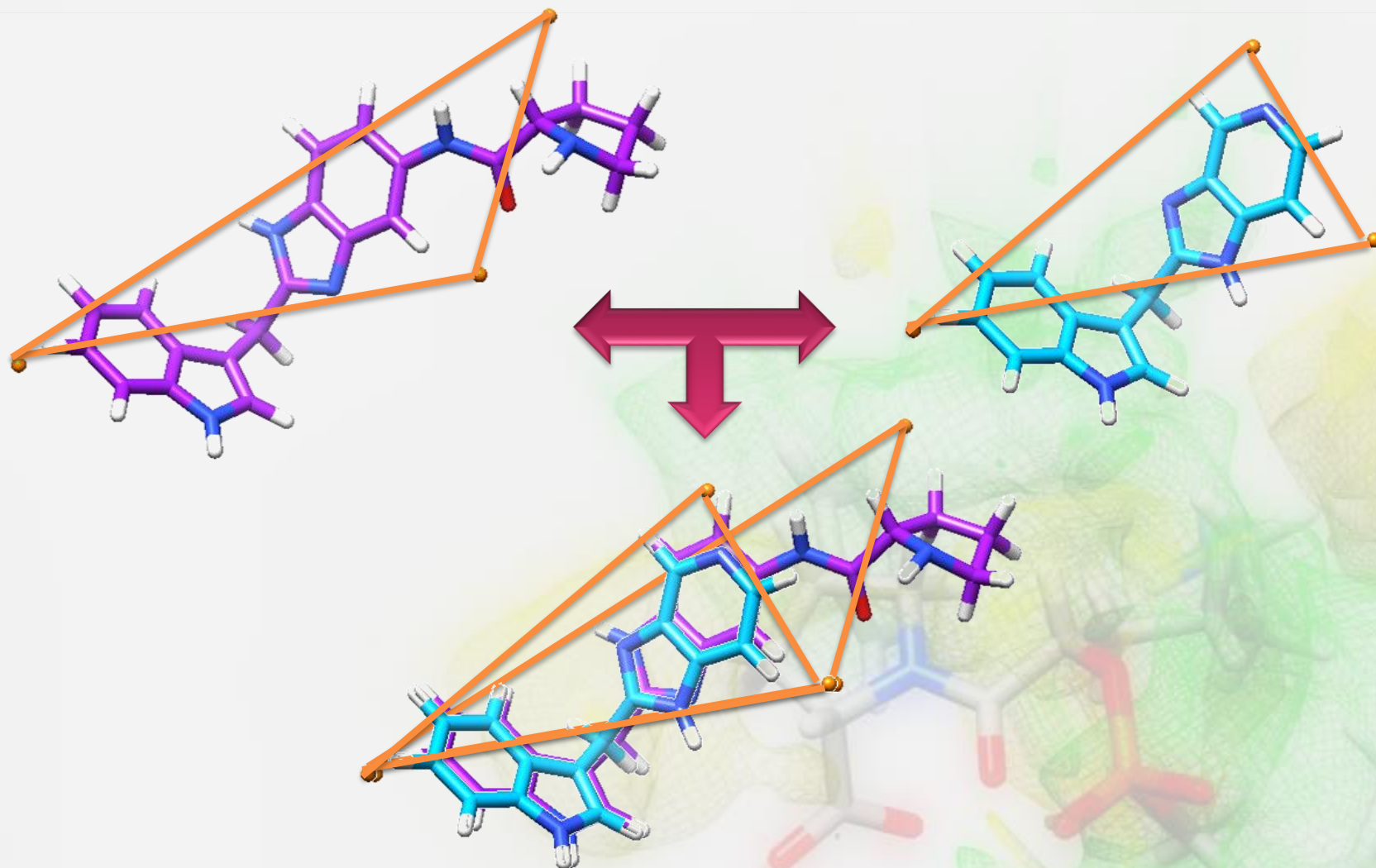
➤ Distanza del punto osservazionale dagli atomi accettori di legami idrogeno

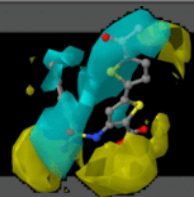
➤ Distanza del punto osservazionale dagli atomi donatori di legami idrogeno



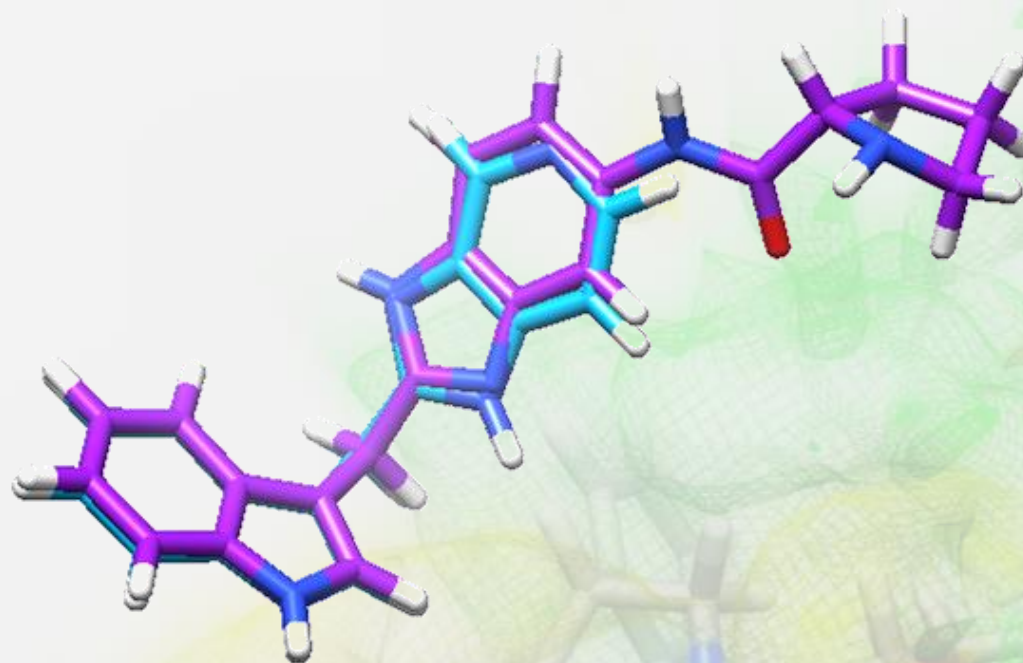


PROGRAMMI DI SOVRAPPOSIZIONE MOLECOLARE : SURFLEX





PROGRAMMI DI SOVRAPPOSIZIONE MOLECOLARE : SURFLEX

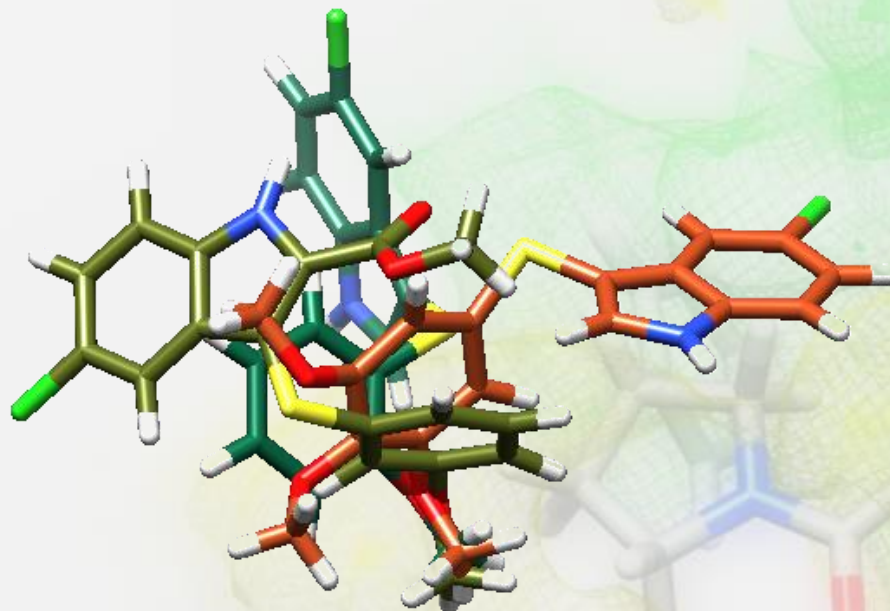




PROGRAMMI DI SOVRAPPOSIZIONE MOLECOLARE : SURFLEX

Limiti:

- ❖ *Molecole devono essere strutturalmente simili;*
- ❖ *Conformazione da utilizzare;*

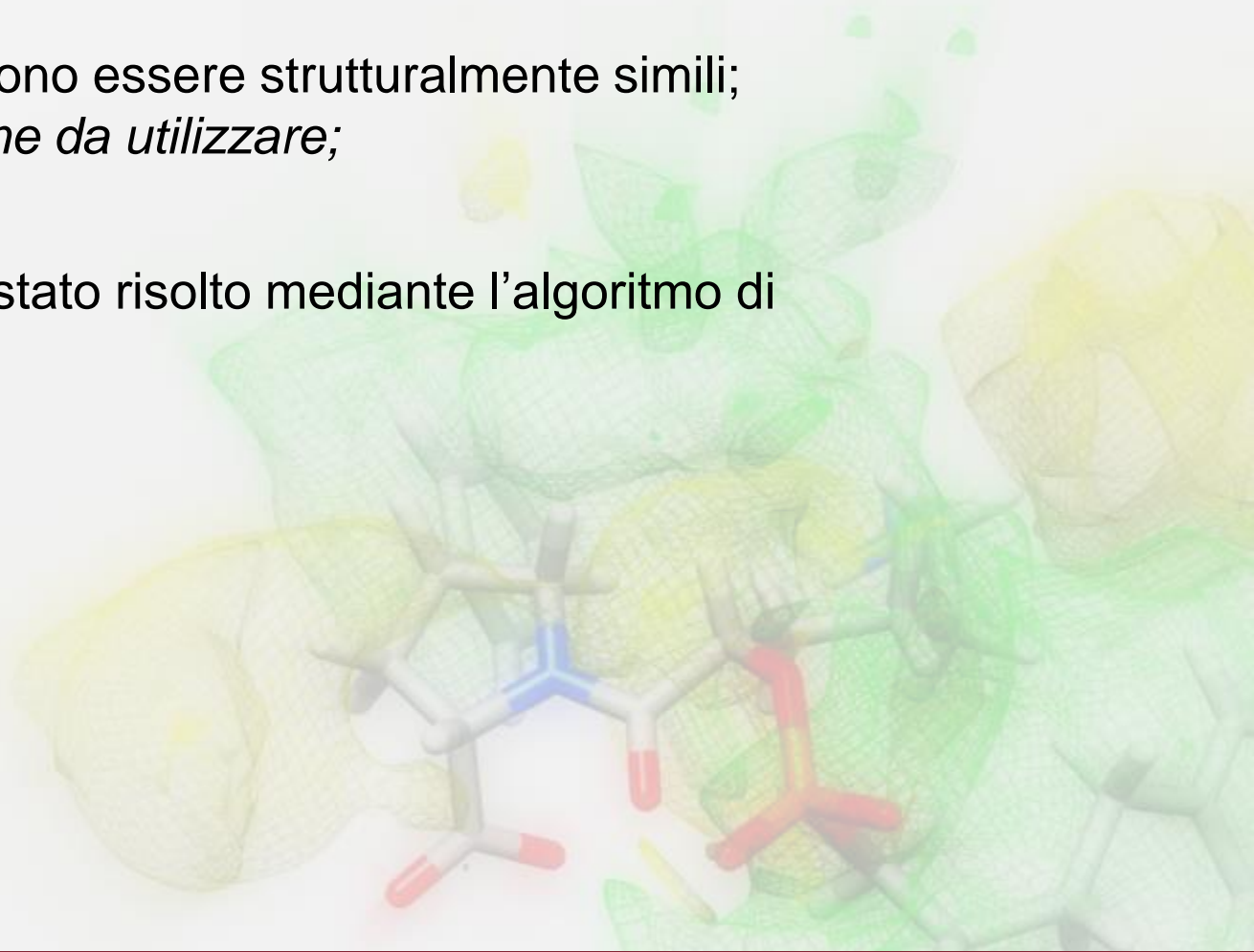


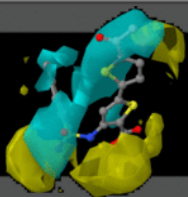


Limiti:

- ❖ Molecole devono essere strutturalmente simili;
- ❖ *Conformazione da utilizzare;*

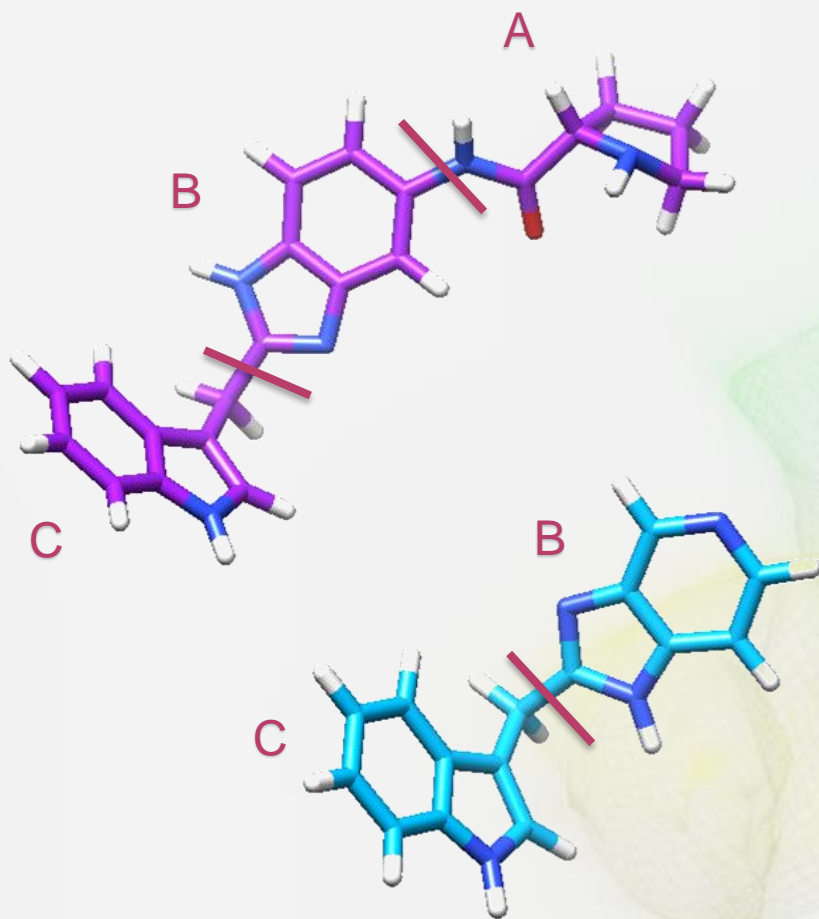
Questo problema è stato risolto mediante l'algoritmo di **Hammerhead**.





PROGRAMMI DI SOVRAPPOSIZIONE MOLECOLARE : SURFLEX

L'algoritmo di **Hammerhead** prevede alcuni step:



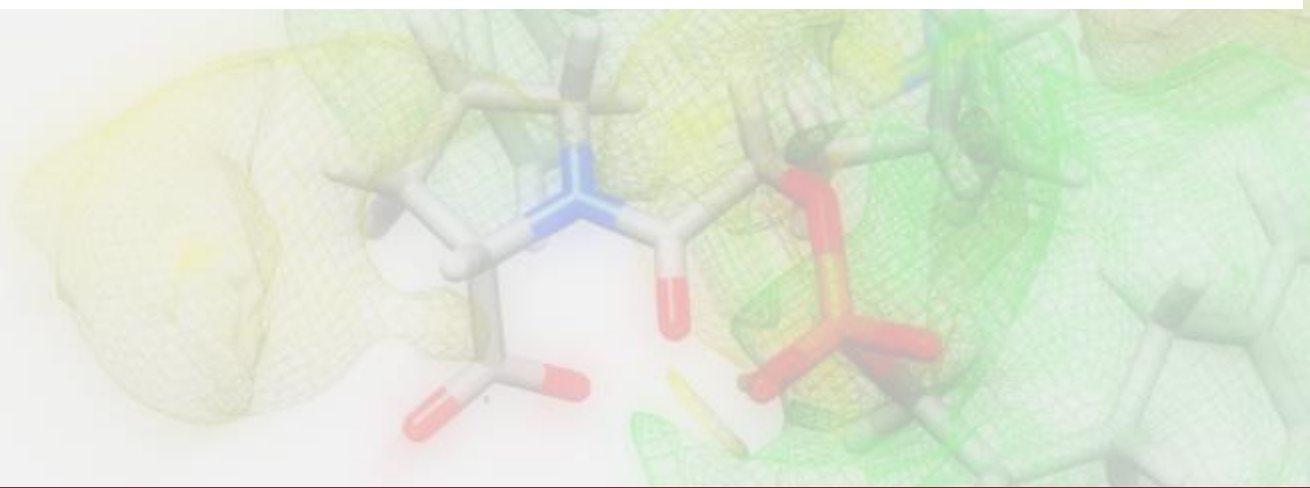
- Frammenti non devono contenere meno di 6 atomi;
- Ciascun frammento viene analizzato conformazionalmente, da default vengono trovati i primi 10 migliori conformeri migliori;
- Ciascun conformero di ciascun frammento è allineato sulla struttura di riferimento, massimizzando la similarità morfologica mediante un processo fattoriale;

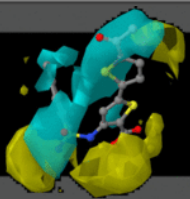


Morphological similarity: A 3D molecular similarity method correlated with protein-ligand recognition

Ajay N. Jain

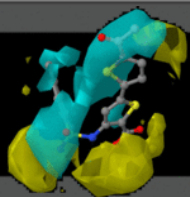
UCSF Cancer Center, Box 0128, San Francisco, CA 94143-0128, U.S.A. E-mail: ajain@cc.ucsf.edu



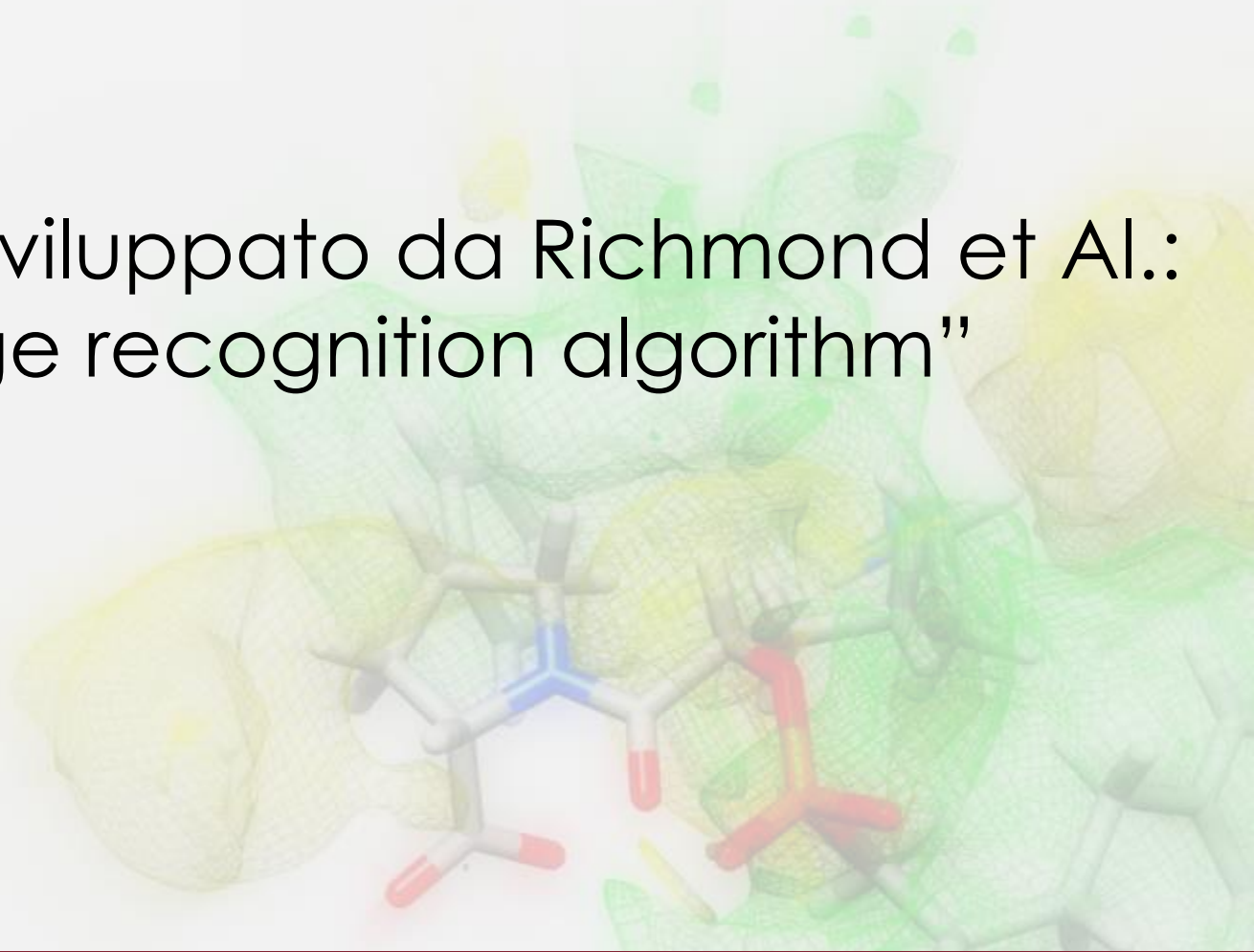


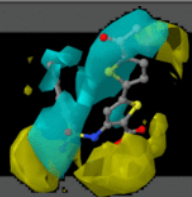
I Criteri mediante il quale l'allineamento può essere sviluppato sono diversi:

- Similarità morfologica;
 - **Atomico**;
 - Farmacoforico;
- campi del potenziale elettrostatico molecolare (MEP);



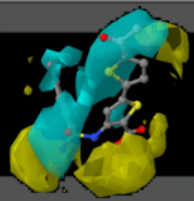
- Algoritmo sviluppato da Richmond et Al.:
“image recognition algorithm”





LAP: LINEAR ASSIGNMENT PROBLEM

		COMPUTER			
		1	2	3	4
LAVORO	1	1	6	2	5
	2	2	1	0	4
	3	1	3	2	3
	4	0	5	1	6

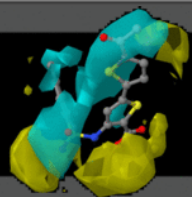


LAP: LINEAR ASSIGNMENT PROBLEM

COMPUTER

LAVORO

	1	2	3	4
1	1	6	2	5
2	2	1	0	4
3	1	3	2	3
4	0	5	1	6



LAP: LINEAR ASSEGNMENT PROBLEM

L
A
V
O
R
O

COMPUTER

	1	2	3	4
1	1	6	2	5
2	2	1	0	4
3	1	3	2	3
4	0	5	1	6

Lavoro totale = 1 + 1 + 3 + 1

= 6

Costo totale più basso
ottenibile



LAP: LINEAR ASSEGNMENT PROBLEM

L
A
V
O
R
O

1
2
3
4

COMPUTER

	1	2	3	4
1	1	6	2	5
2	2	1	0	4
3	1	3	2	3
4	0	5	1	6

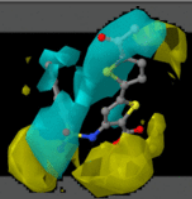
Lavoro totale = 1 + 1 + 3 + 1

= 6

Costo totale più basso
ottenibile

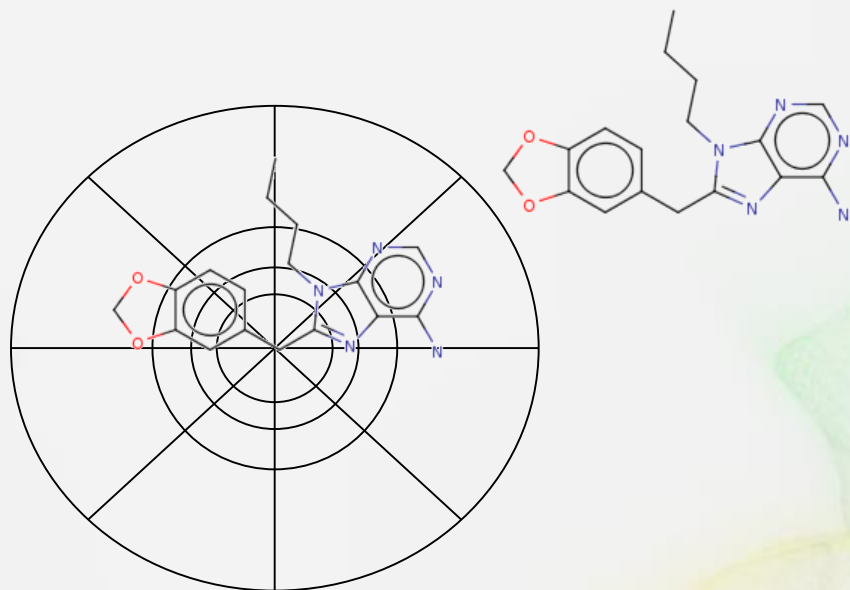


Algoritmo di Richmond

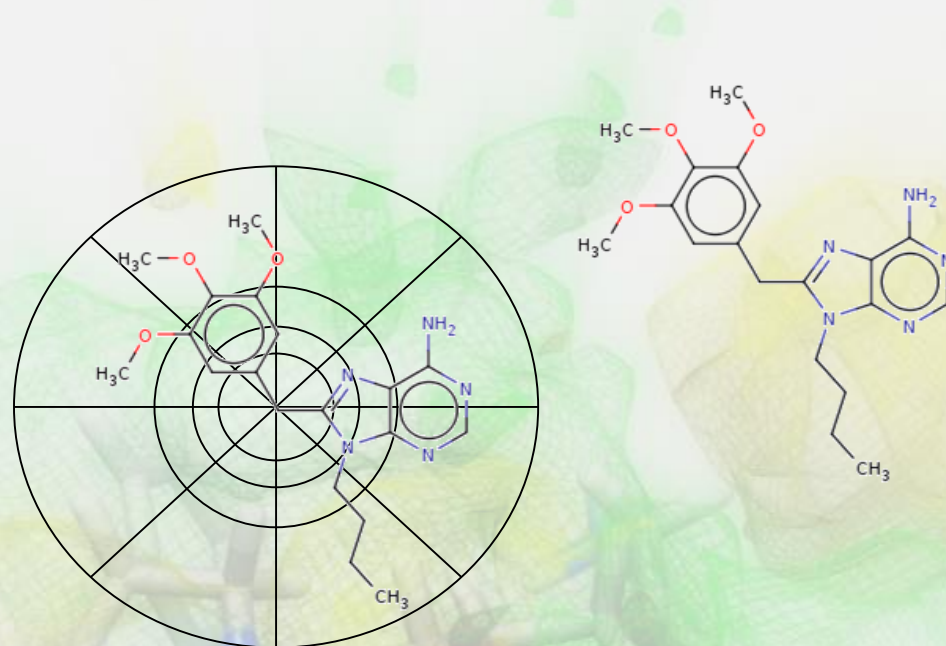


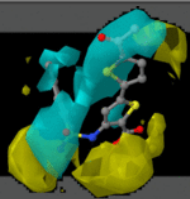
OPEN3DALIGN : ATOM-BASED

Molecola A



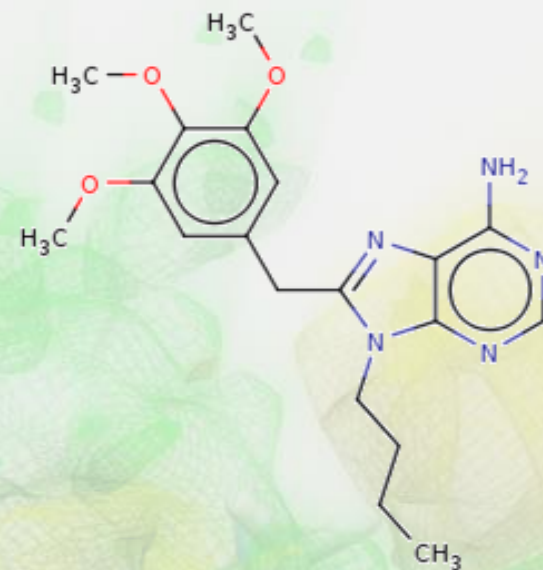
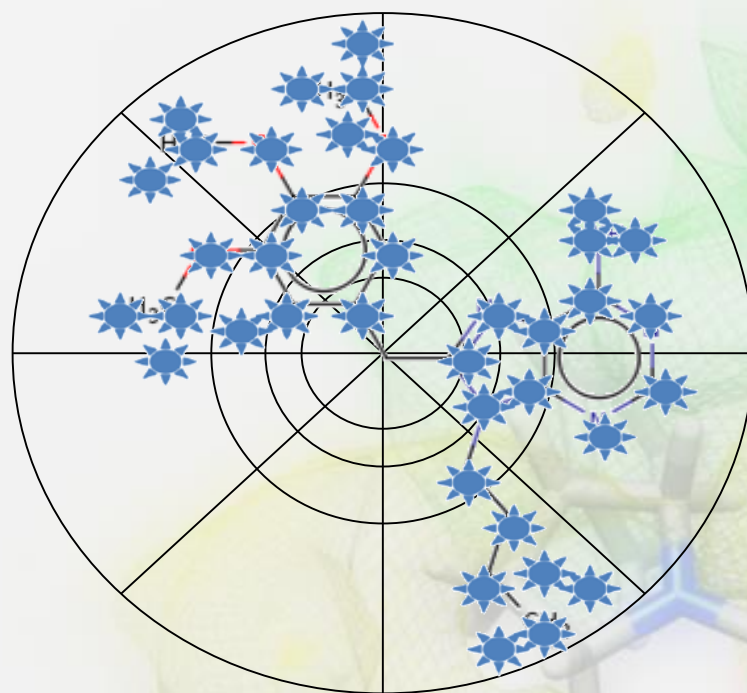
Molecola B

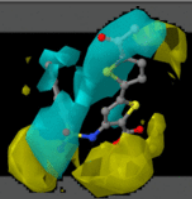




OPEN3DALIGN : ATOM-BASED

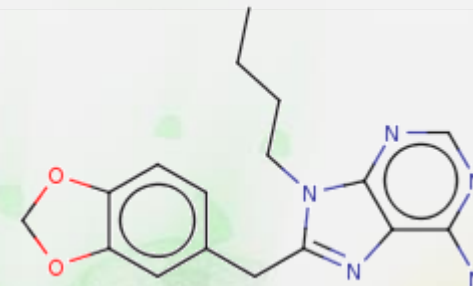
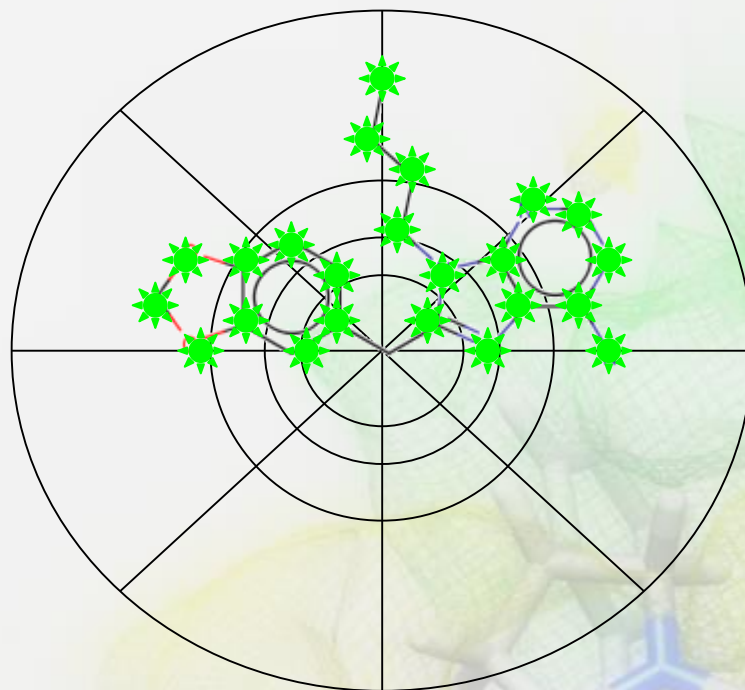
Molecola A





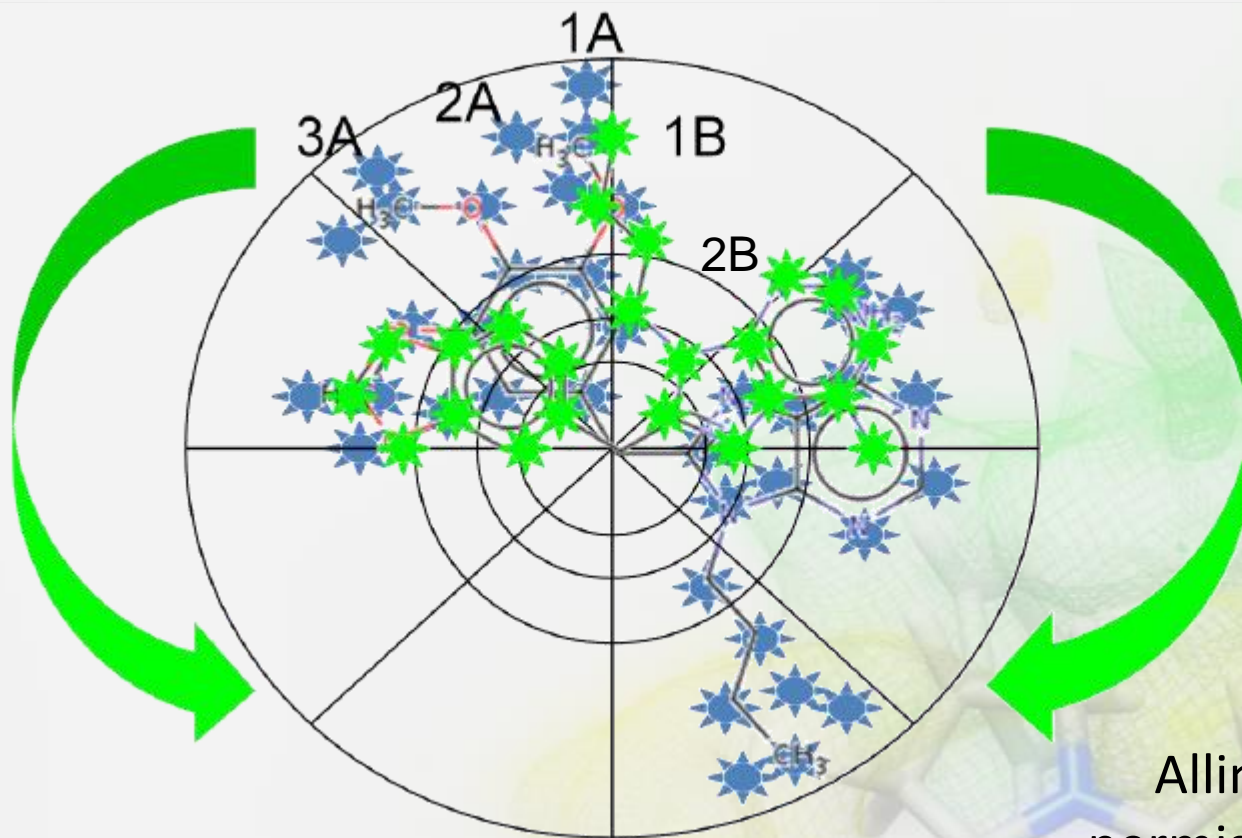
OPEN3DALIGN : ATOM-BASED

Molecola B





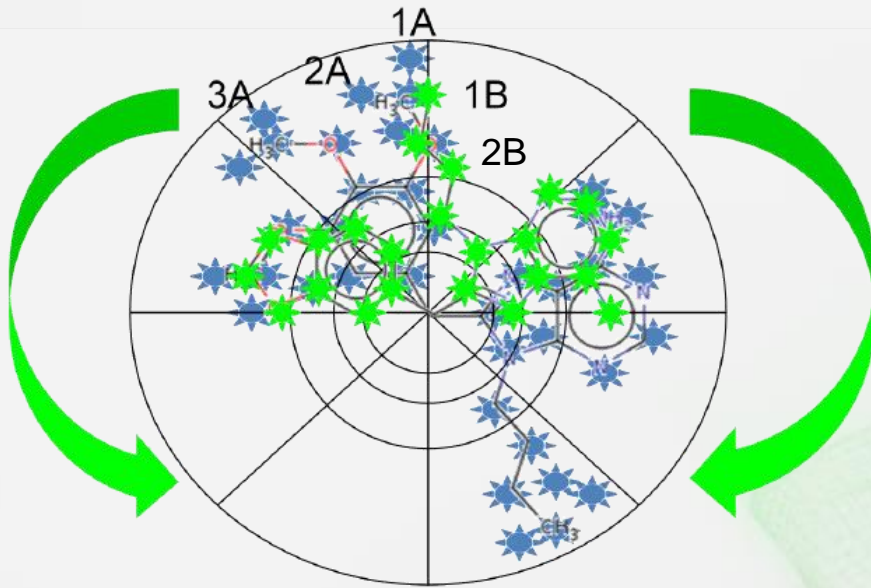
OPEN3DALIGN : ATOM-BASED



Allineamento rigido:
permissione di movimenti
isometrici: traslazionali e
rotazionali



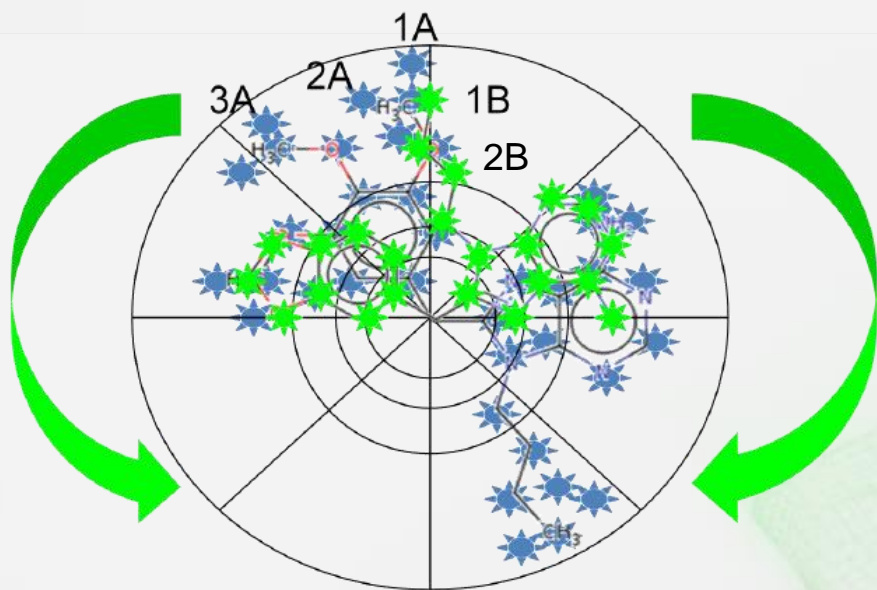
OPEN3DALIGN : ATOM-BASED



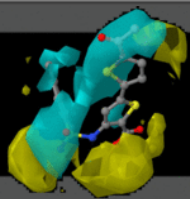
	A1	A2	A3	A4
B1	2	4		
B2	0	9		
B3	5			
B4	3			



OPEN3DALIGN : ATOM-BASED



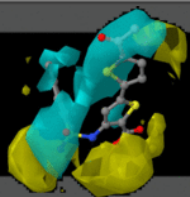
	A1	A2	A3	A4
B1	2	4	4	8
B2	0	9	1	3
B3	5	5	7	2
B4	3	4	1	5



	A1	A2	A3	A4
B1	2	4	4	8
B2	0	9	1	3
B3	5	5	7	2
B4	3	4	1	5

ALGORITMO DI RICHARD ET AL.

Ci aiuta a individuare le sovrapposizioni dei vari punti dei due istogrammi, che richiedono il minor lavoro possibile.



COSTO DI ALLINEAMENTO E LAP

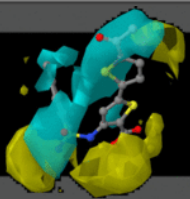
	A1	A2	A3	A4
B1	2	4	4	8
B2	0	9	1	3
B3	5	5	7	2
B4	3	4	1	5

Costo minore possibile = 9

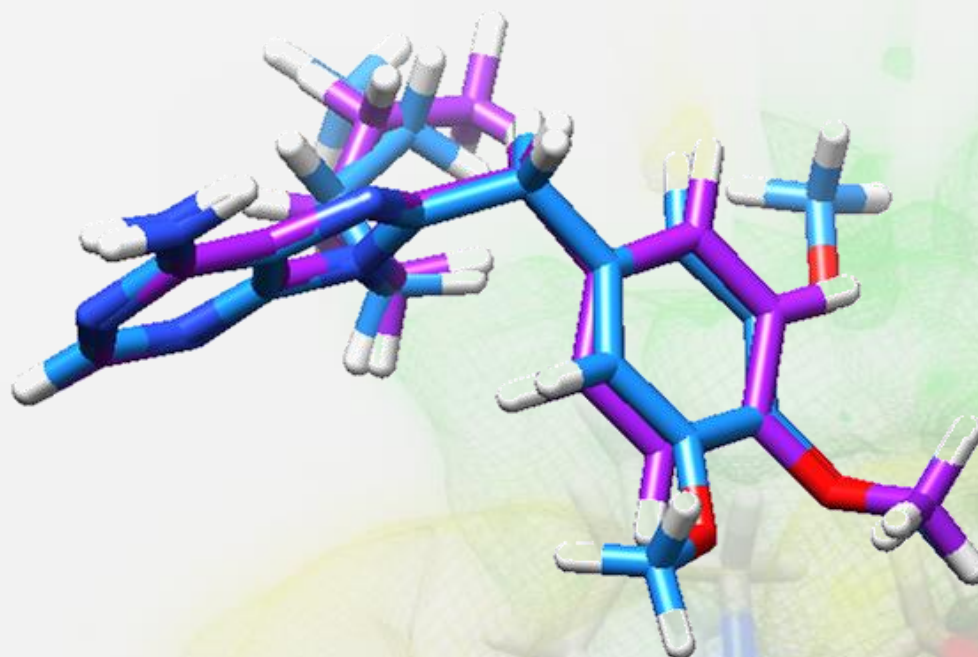
SCORES

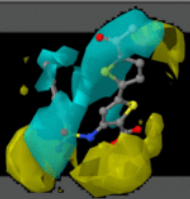
ALGORITMO DI RICHARD ET AL.

Ci aiuta a individuare le sovrapposizioni dei vari punti dei due istogrammi, che richiedono il minor lavoro possibile.



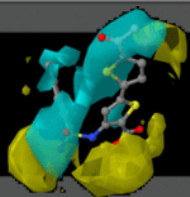
PROGRAMMI DI SOVRAPPOSIZIONE MOLECOLARE



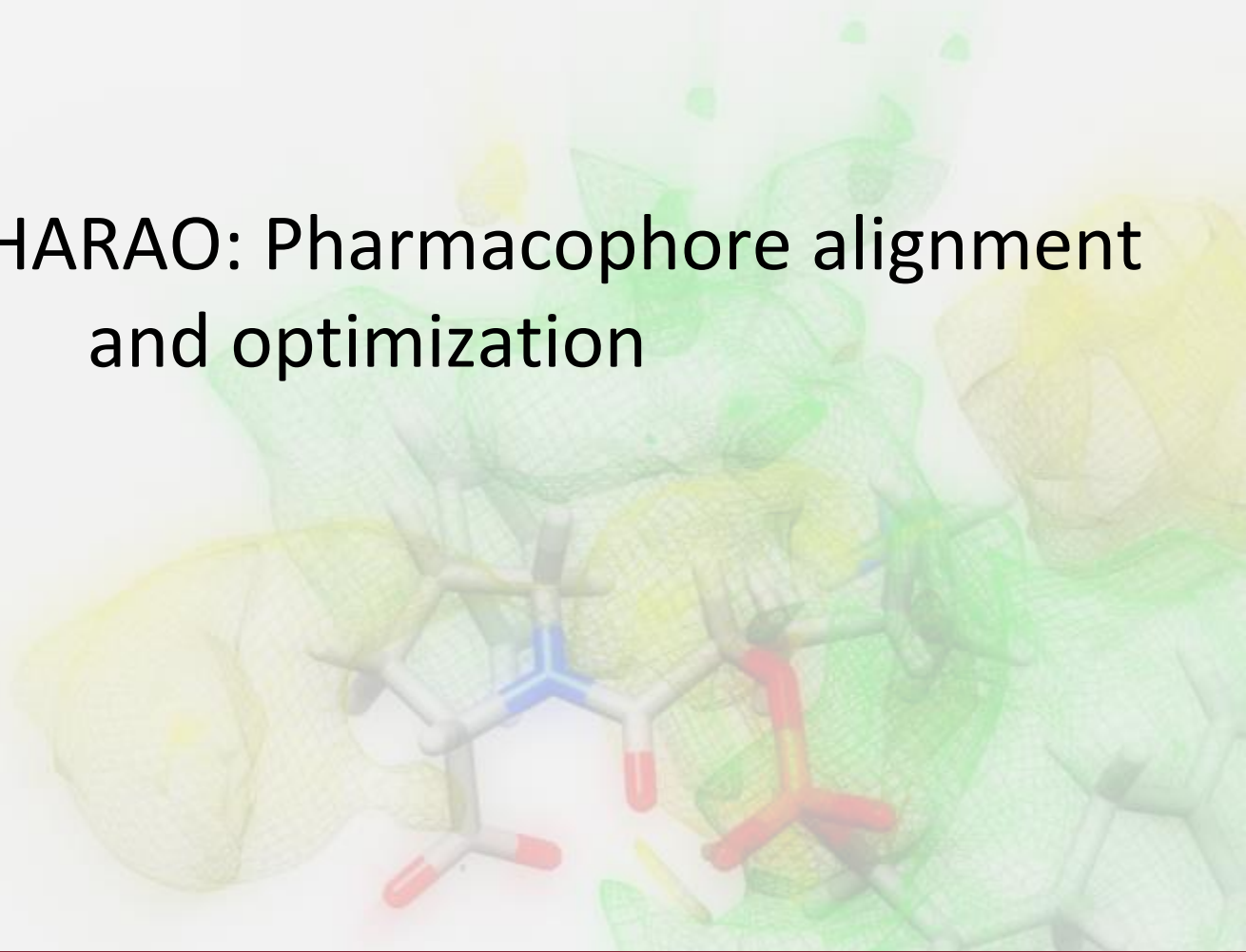


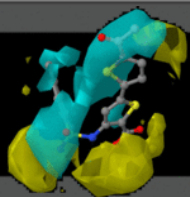
I Criteri mediante il quale l'allineamento può essere sviluppato sono diversi:

- Similarità morfologica;
 - Atomico;
 - ***Farmacoforico;***
- campi del potenziale elettrostatico molecolare (MEP)

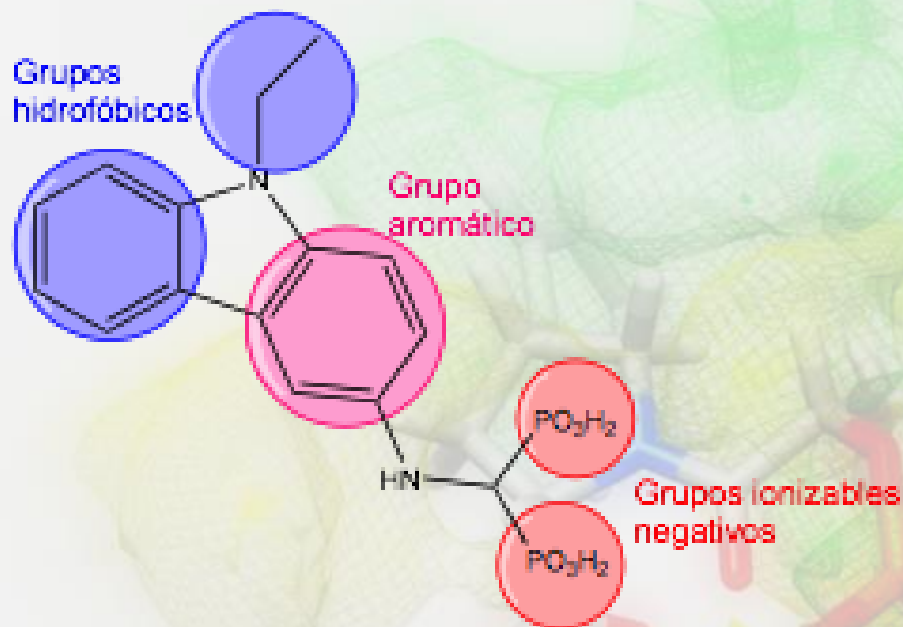


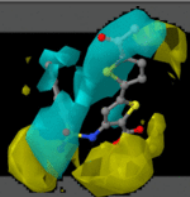
- Algoritmo PHARAO: Pharmacophore alignment and optimization





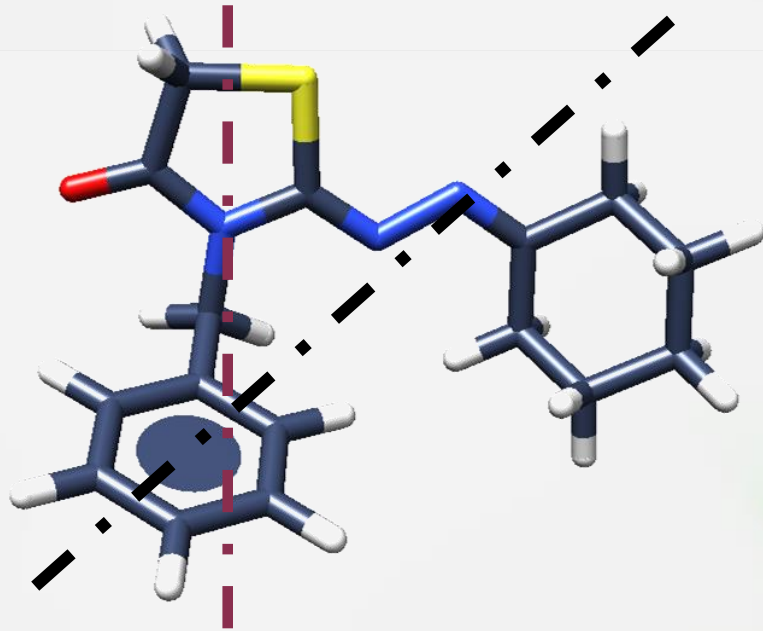
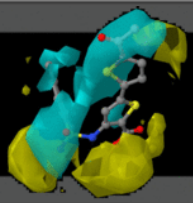
- Farmacoforo: Insieme di caratteri chimico-fisici (detti anche **elementi farmacoforici**) che in opportune orientazioni spaziali tridimensionali sono responsabili dell'interazione farmaco-recettore, e quindi dell'attività biologica di una molecola.



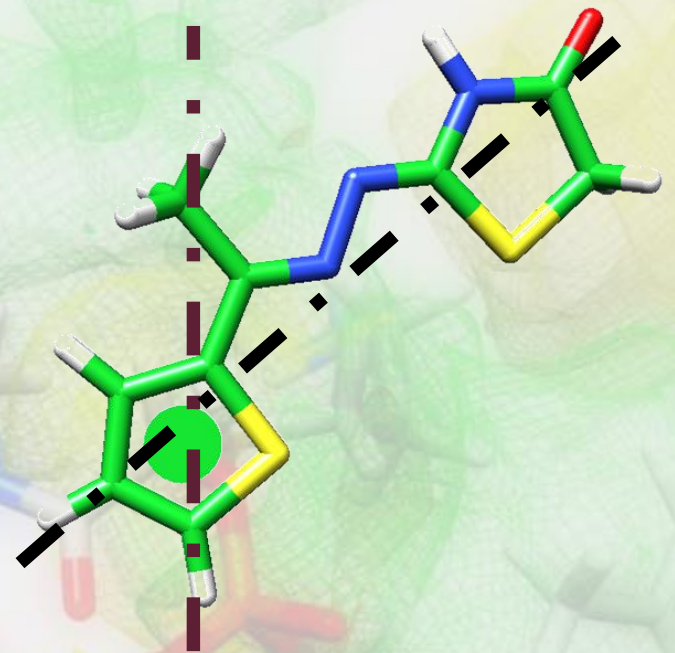


- Primo step: l'individuazione del gruppo farmacoforico.

Gruppo funzionale	descrizione
Anello aromatico	Struttura ciclica planare in cui tutti gli atomi coinvolti condividono tramite i loro orbitali p un totale di $4n+2$ elettroni, dove n è un intero positivo
Donatori legami idrogeno	-Azoto o Ossigeno -Carica formale non negativa -Almeno un H legato covalentemente
Accettori legami idrogeno	-Azoto o Ossigeno -Carica formale non positiva -Almeno un l'on pair disponibile -Accessibilità atomo
Centri di carica	Punto della molecola in cui una carica si è delocalizzata

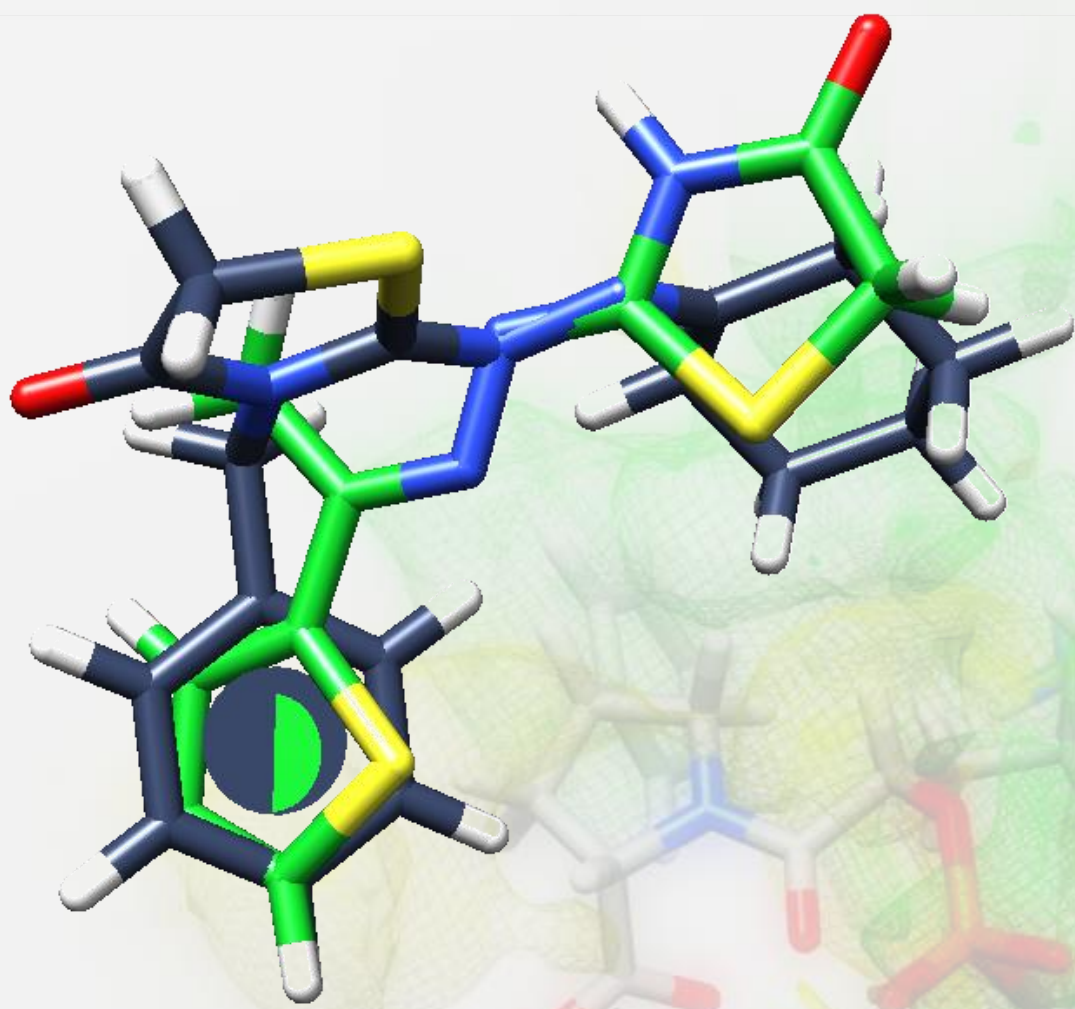


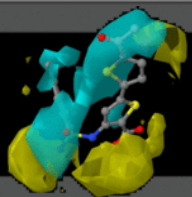
Si individuano gli assi che passano per il numero maggiore possibile di punti farmacoforici individuati nelle molecole da allineare e si passa ad allineare questi assi osservati.





OPEN3DALIGN : PHARM-BASED





TANIMOTO

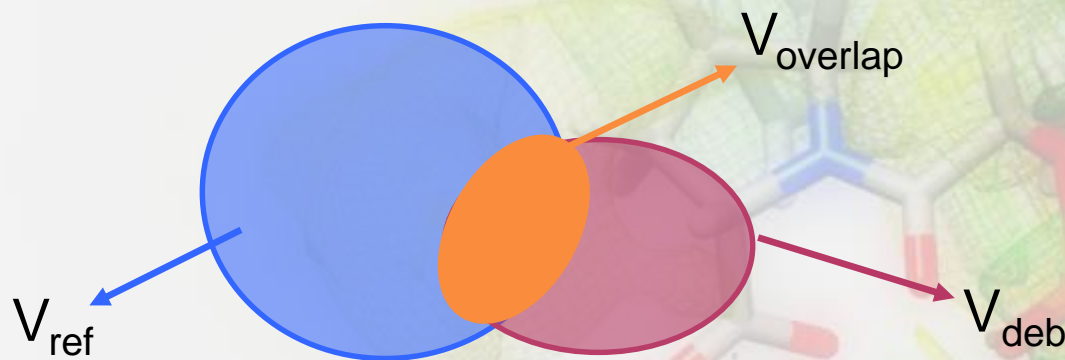
$$= \frac{V_{OVERLAP}}{V_{REF} + V_{DB} - V_{OVERLAP}}$$

$$TVERSKY_REF = \frac{V_{OVERLAP}}{V_{REF}}$$

TVERSKY

$$_DB = \frac{V_{OVERLAP}}{V_{DB}}$$

Vref = volume del primo farmacofooro
Vdb=volume del secondo frmacofooro





I software che usano come criteri di allineamento quello atomico e quello farmacoforico non sono dotati di uno algoritmo interno per la ricerca dei conformeri ottimali da usare nella fase di allineamento.

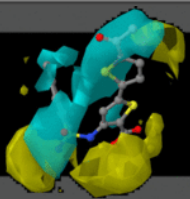
Per questo vengono usati programmi esterni:

- ✓ Balloon
- ✓ QMD



PROGRAMMI DI SOVRAPPOSIZIONE MOLECOLARE

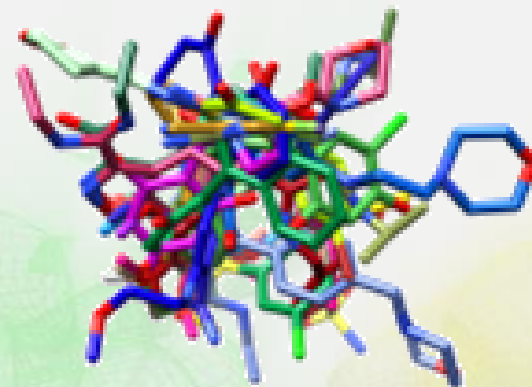
Principio d'allineamento	Software
Similarita morfologica	Suflex-sim
Atom-based	OpenAlign Align-it Kcombu
Pharm-based	OpenAlign Shape-it
MEP	Shaep



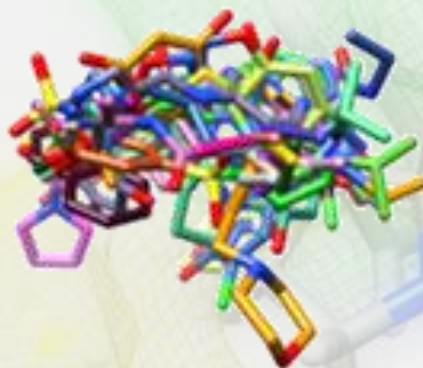
PROGRAMMI DI SOVRAPPOSIZIONE MOLECOLARE



Software :Surflex-sim
Referenza: meno attiva



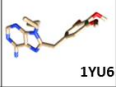
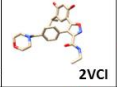

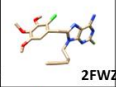
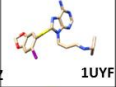
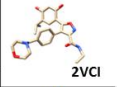





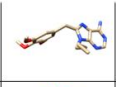
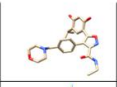
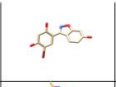
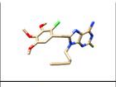
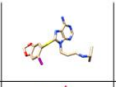
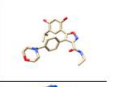

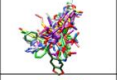
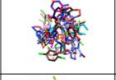
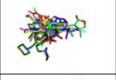
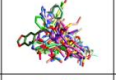
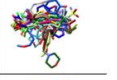
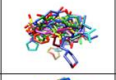
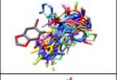
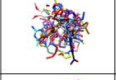
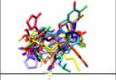
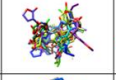
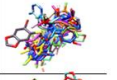
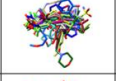
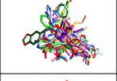
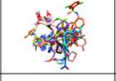
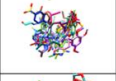
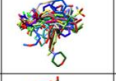
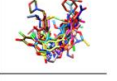
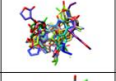
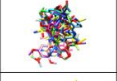
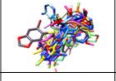
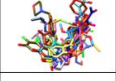
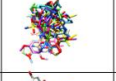
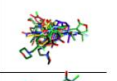
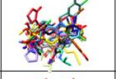
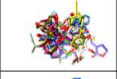
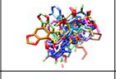
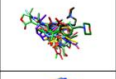
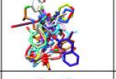
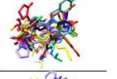
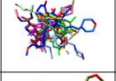
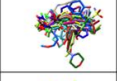
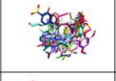
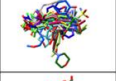
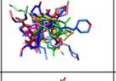
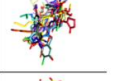
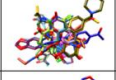
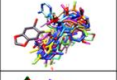
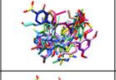
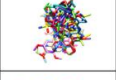
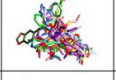
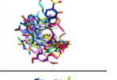
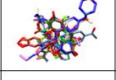
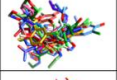
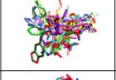
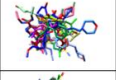
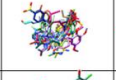
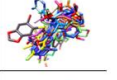
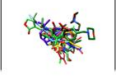

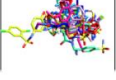


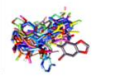
Software :Open3Dalign
pharmaphoric
Referenza: meno attiva



Software :Open3Dalign
atomic
Referenza: meno attiva



PROGRAMMI DI SOVRAPPOSIZIONE MOLECOLARE

		-	+	-	+	+	+	
		ACTIVE	ACTIVE	FLEXIBLE	FLEXIBLE	HEAVY	VOL	
SURFLEX	REF	 1YU6	 2VCI	 1BM9	 2FWZ	 1UYF	 2VCI	
	SURFLEX							
Open3DALIGN (atom-based)	REF							
	CRY	ATOM						
		MIXED						
		PHARM						
	GBSA	ATOM						
		MIXED						
		PHARM						
	VACUUM	ATOM						
		MIXED						
		PHARM						





COSTRUZIONE DI UN MODELLO

Scelta del
training set

Allineamento
delle
molecole

Calcolo del
campi di
interazione
molecolare
(MIF)

Generazione
del modello
statistico

Validazione
interna,
esterna e
y-scrambling