

L'allineamento nel molecular modelling e cenni di 3-D QSAR



SAPIENZA
UNIVERSITÀ DI ROMA

Manuela Sabatino

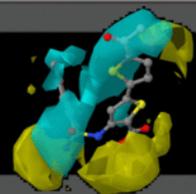
Martedì
24 Maggio 2016



(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

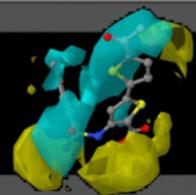
Scelta del training set

Allineamento delle molecole

Calcolo dei campi di interazione molecolare (MIF)

Generazione del modello statistico

Validazione interna, esterna e y-scrambling



COSTRUZIONE DI UN MODELLO

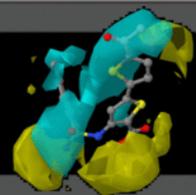
Scelta del training set

Allineamento delle molecole

Calcolo del campi di interazione molecolare (MIF)

Generazione del modello statistico

Validazione interna, esterna e y-scrambling



COSTRUZIONE DI UN MODELLO

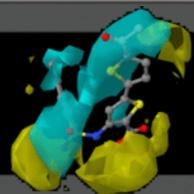
Scelta del
training set

Allineamento
delle
molecole

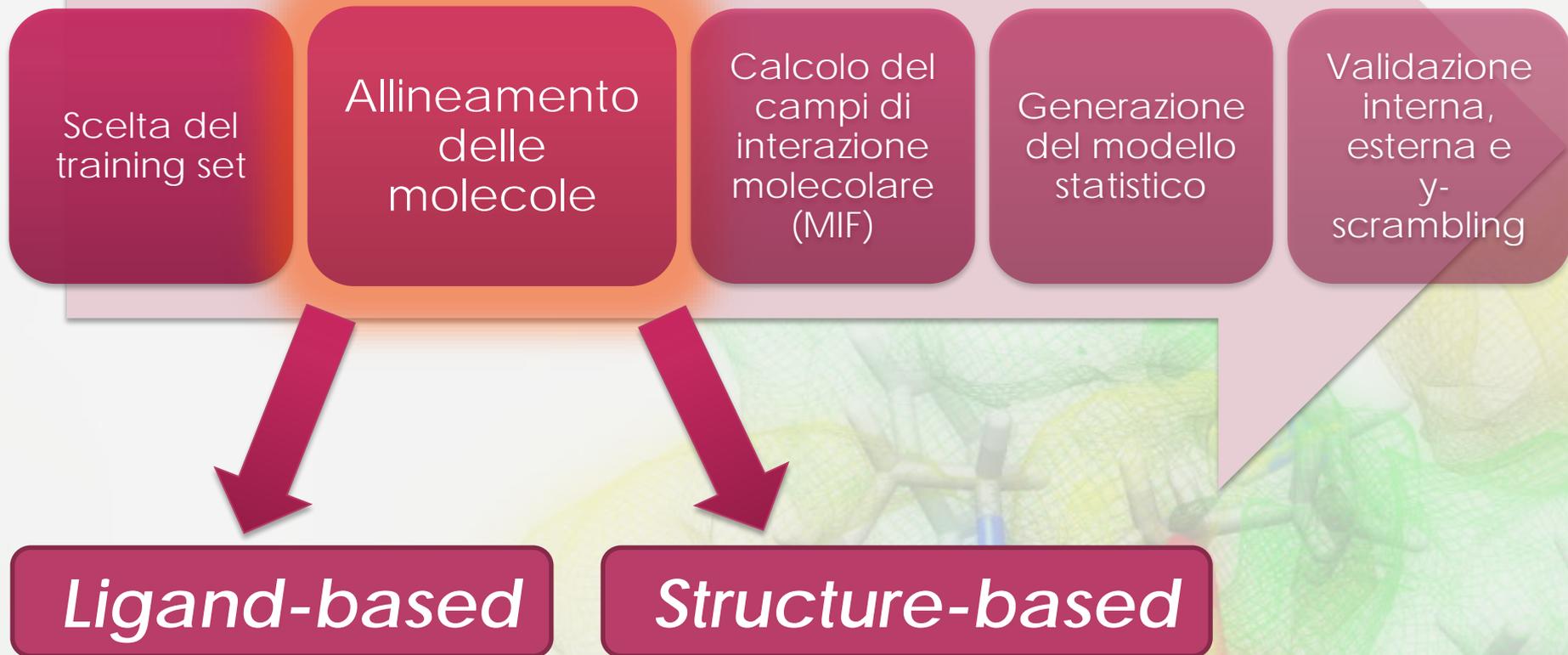
Calcolo del
campi di
interazione
molecolare
(MIF)

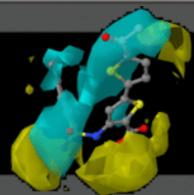
Generazione
del modello
statistico

Validazione
interna,
esterna e
y-
scrambling



TIPOLOGIE DI ALLINEAMENTO





APPROCCIO STRUCTURE-BASED

rcsb
www.rcsb.it

RCSB PDB
PROTEIN DATA BANK

“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

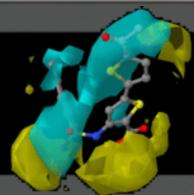
November Molecule of the Month

glutamate

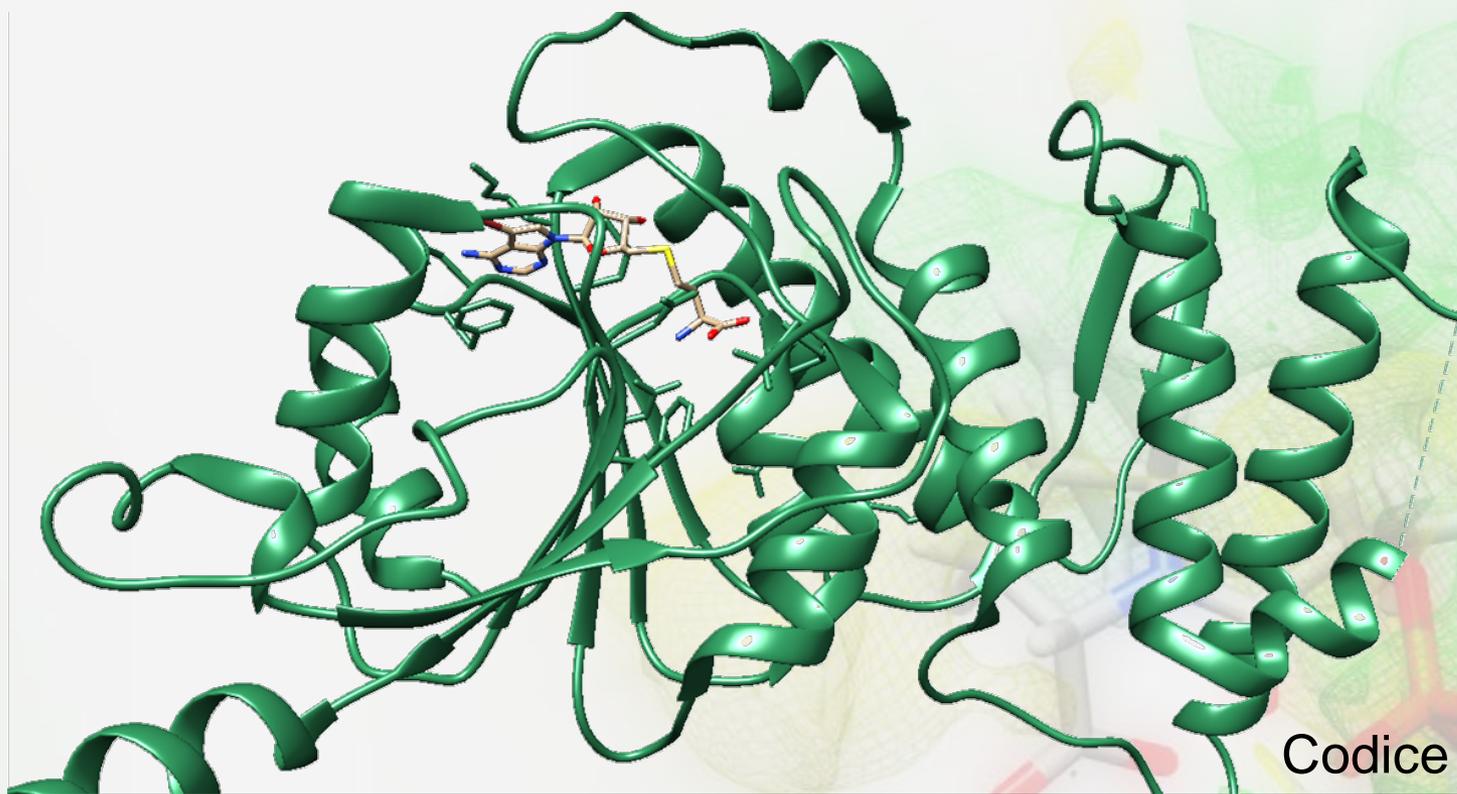
ivermectin

Glutamate-gated Chloride Receptors

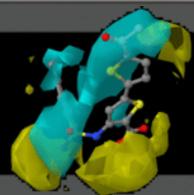
Feedback



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
------	------

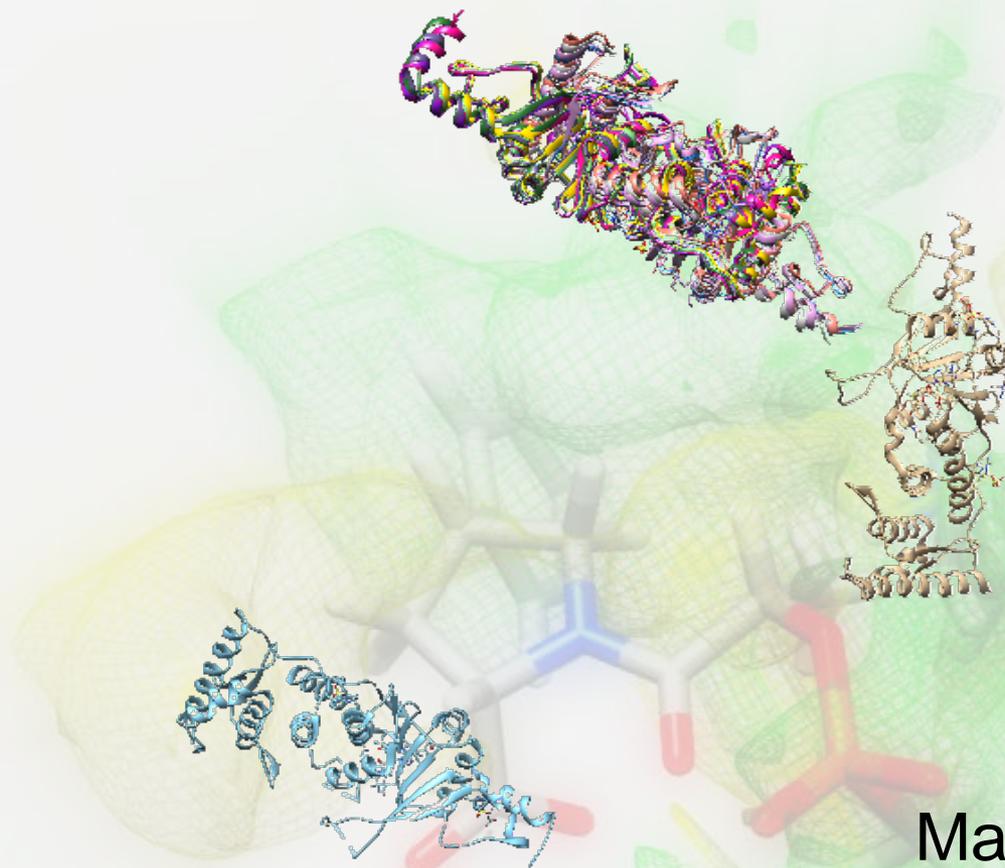
4ER0	2.50
------	------

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
------	------

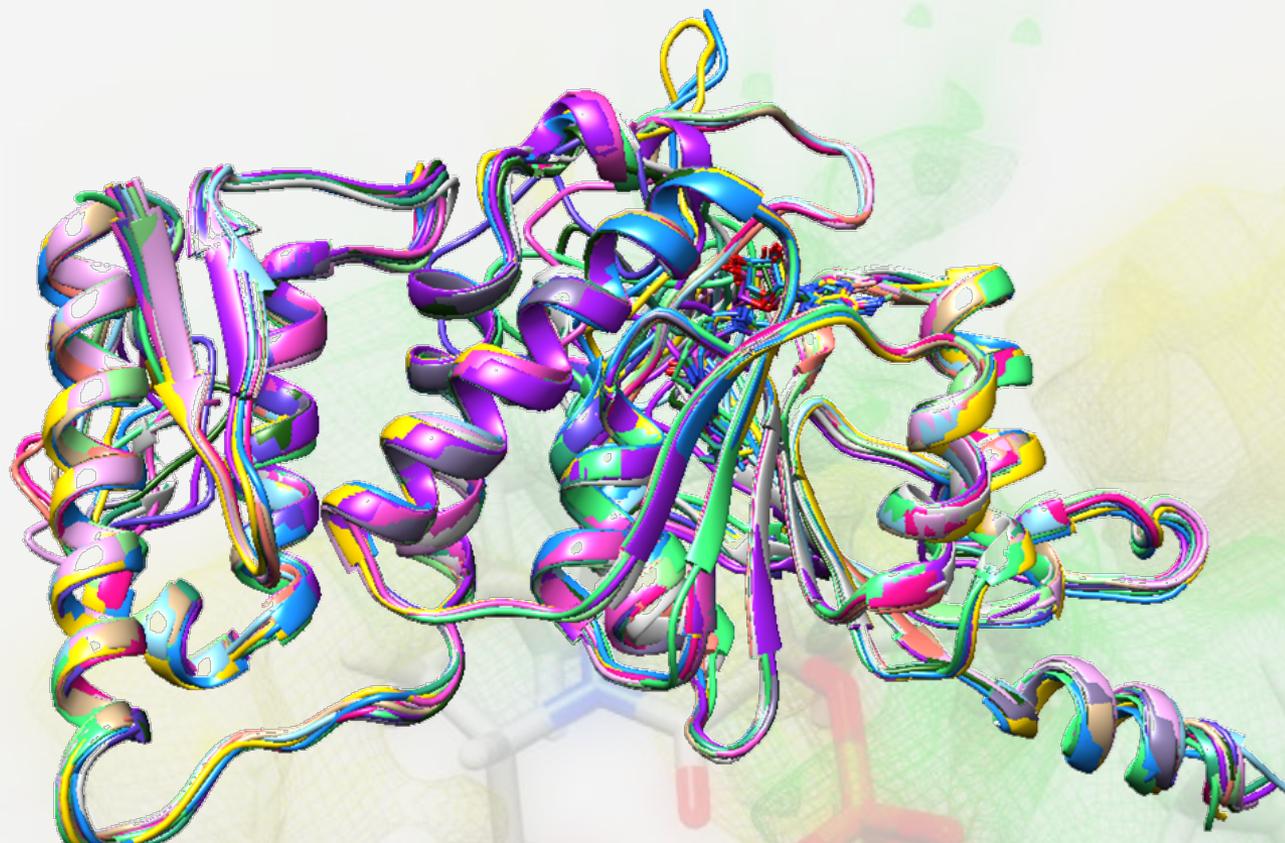
4ER0	2.50
------	------

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

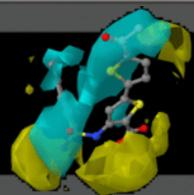
4ER7	2.20
------	------

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

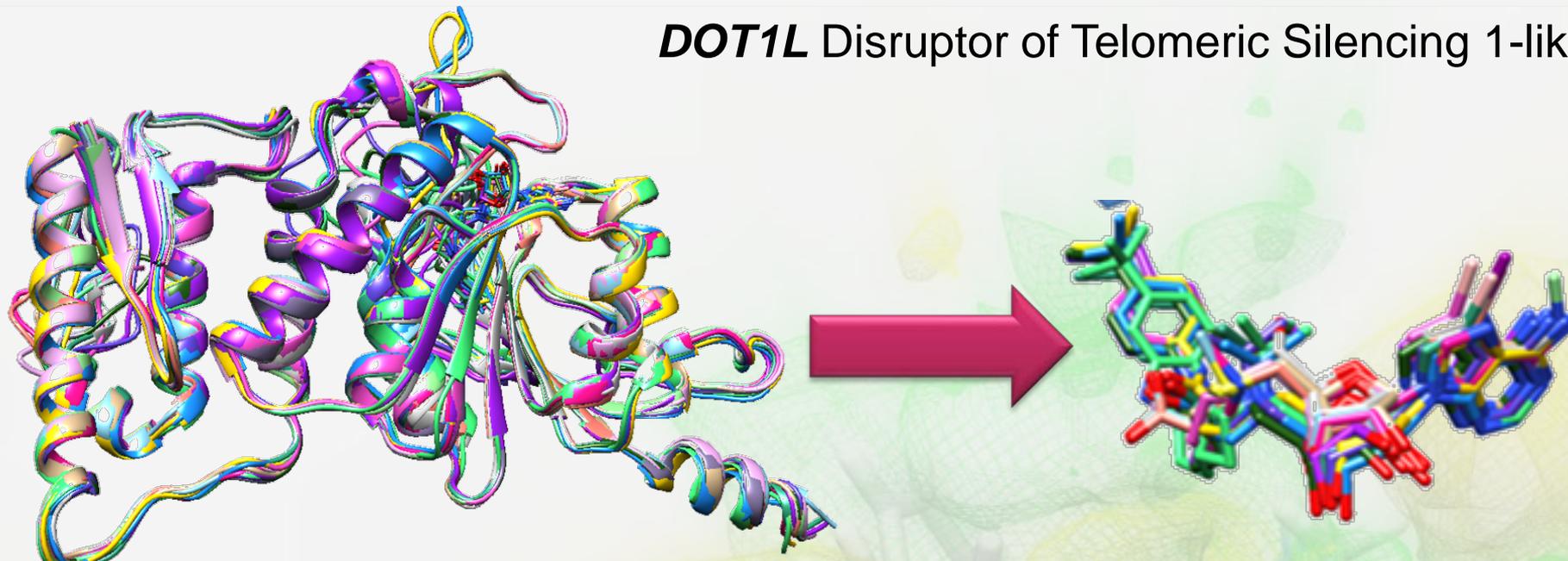
DOT1L Disruptor of Telomeric Silencing 1-like



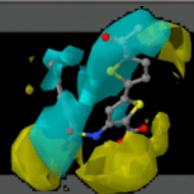
MatchMaker



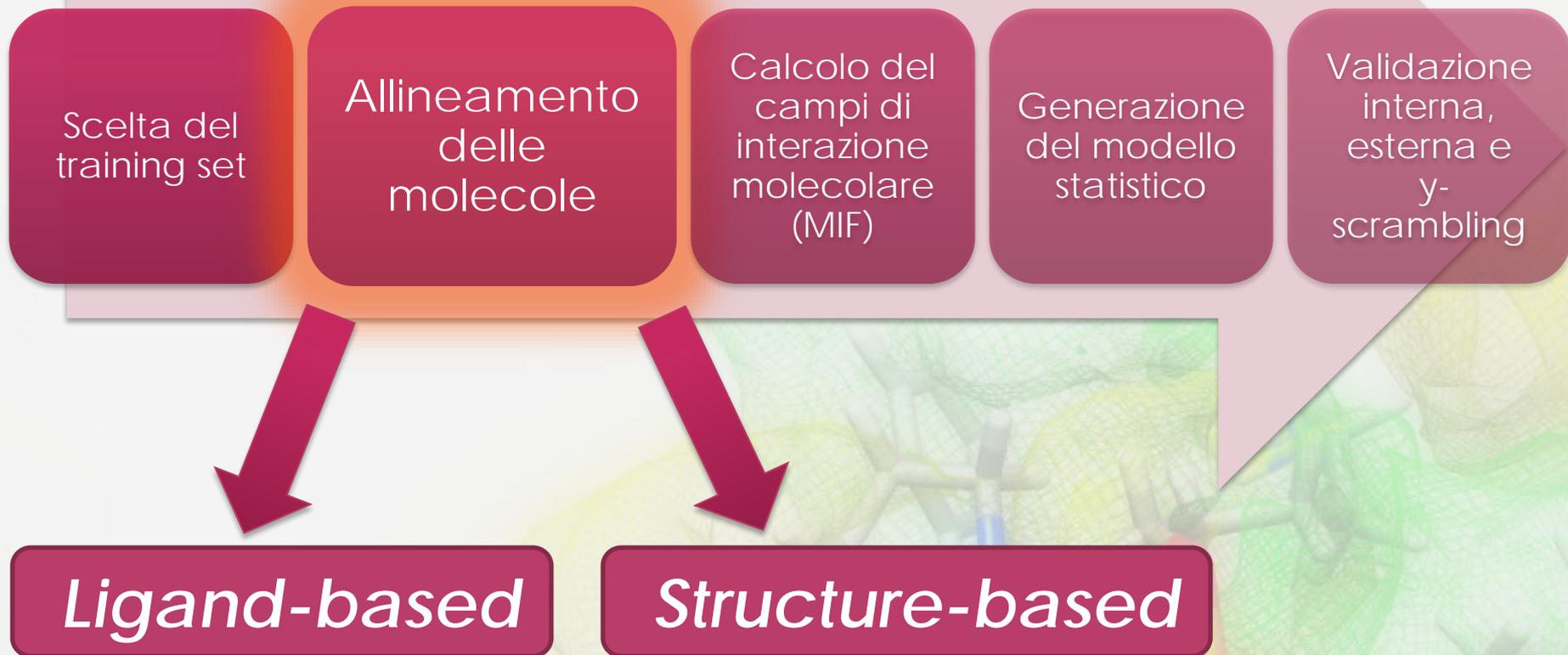
DOT1L Disruptor of Telomeric Silencing 1-like



Allineamento con approccio
STRUCTURE-BASED

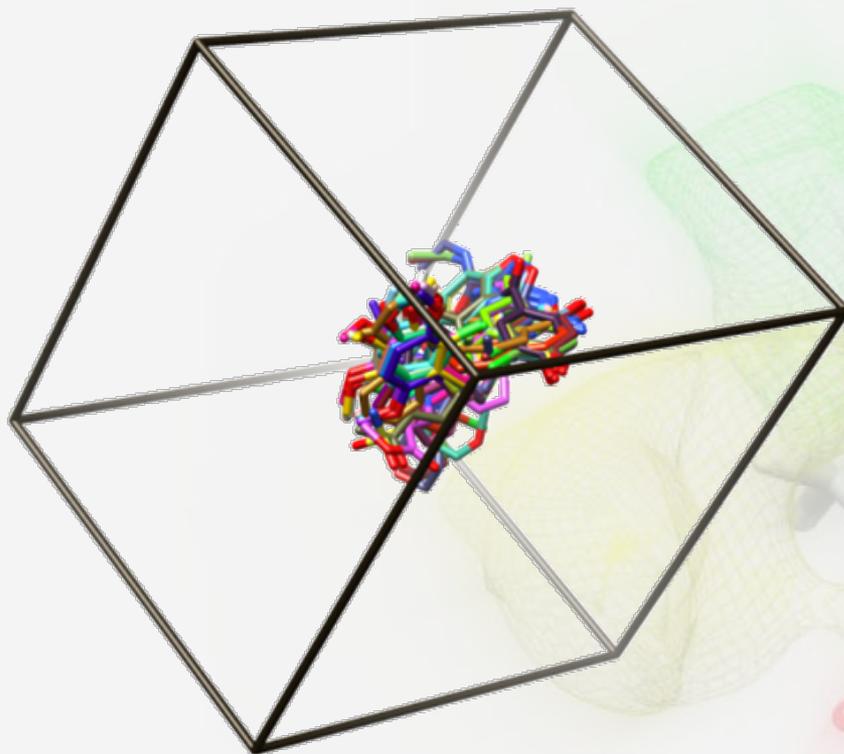


TIPOLOGIE DI ALLINEAMENTO

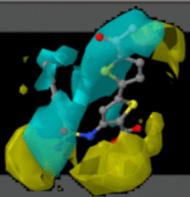




- ✓ Viene utilizzato in assenza di Dati 3-D del target su cui i nostri inibitori agiscono;
- ✓ Utile a individuare una disposizione spaziale del dataset che tenta di mimare una ipotetica binding mode;

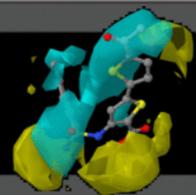


- ✓ Molecola più attiva;
- ✓ Molecola meno attiva;
- ✓ Molecola più flessibile;
- ✓ Molecola meno flessibile;
- ✓ Molecola più pesante;
- ✓ Molecola più lunga e voluminosa;



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

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



COSTRUZIONE DI UN MODELLO

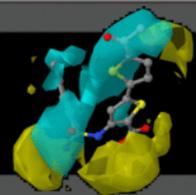
Scelta del training set

Allineamento delle molecole

Calcolo dei campi di interazione molecolare (MIF)

Generazione del modello statistico

Validazione interna, esterna e y-scrambling



COSTRUZIONE DI UN MODELLO

Scelta del training set

Allineamento delle molecole

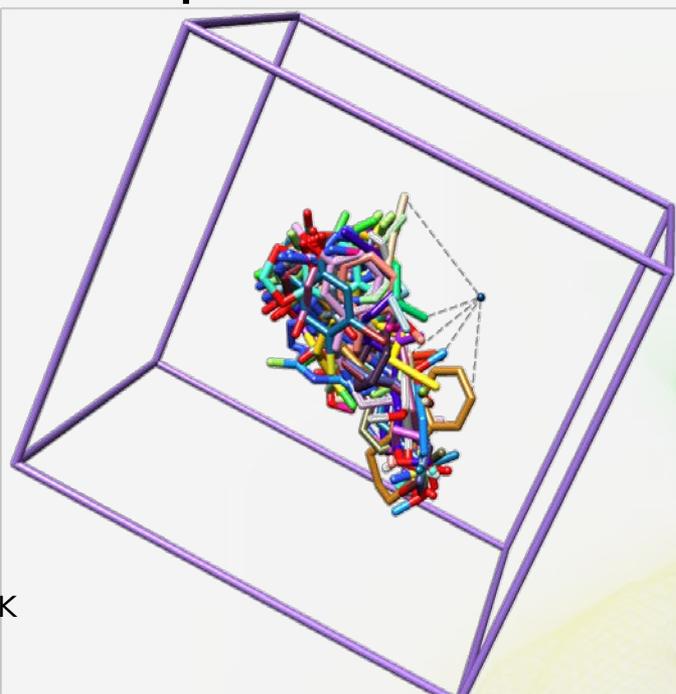
Calcolo del campi di interazione molecolare (MIF)

Generazione del modello statistico

Validazione interna, esterna e y-scrambling



- Campi di interazione molecolare (MIF)



$$F = \frac{q_1 q_2}{r^2} K$$

Dove:

q_1 = La carica del primo atomo

q_2 = La carica del secondo atomo

r = la distanza che intercorre tra i due atomi analizzati

K = la costante di coulomb

I MIF sono le energie che si sviluppano tra le molecole del nostro Training Set e gli atomi Probe, o atomi sonda. Queste energie vengono calcolate mediante le Legge di Coulomb e La Lennard-Jones

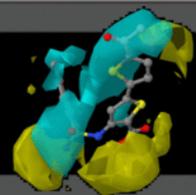
$$V(r) = \frac{A}{r^{12}} - \frac{B}{r^6}$$

$$A = 4\epsilon\delta^{12}$$

$$B = 4\epsilon\delta^6$$

ϵ = Profondità della buca di potenziale

δ = diametro della sfera che approssima l'atomo o la molecola



COSTRUZIONE DI UN MODELLO

Scelta del training set

Allineamento delle molecole

Calcolo dei campi di interazione molecolare (MIF)

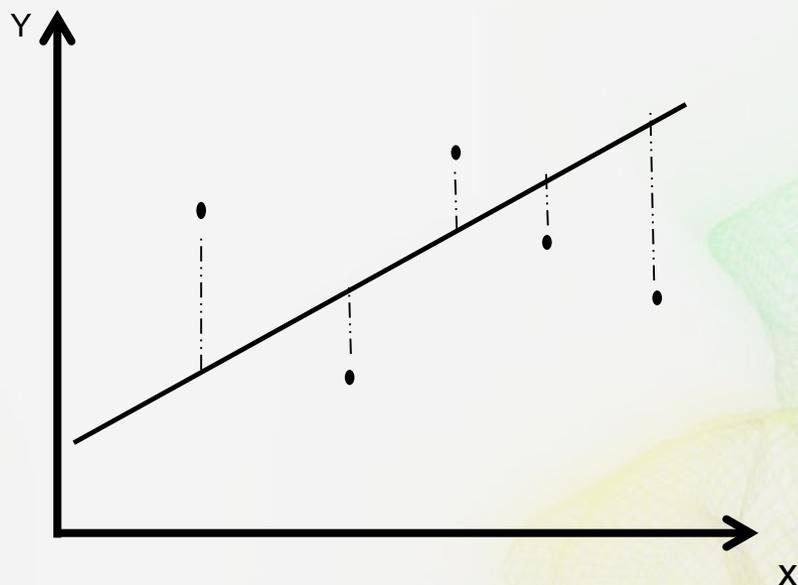
Generazione del modello statistico

Validazione interna, esterna e y-scrambling



- PLS → Partial Least Square;

Tecnica basata sulla teoria dei minimi quadrati.



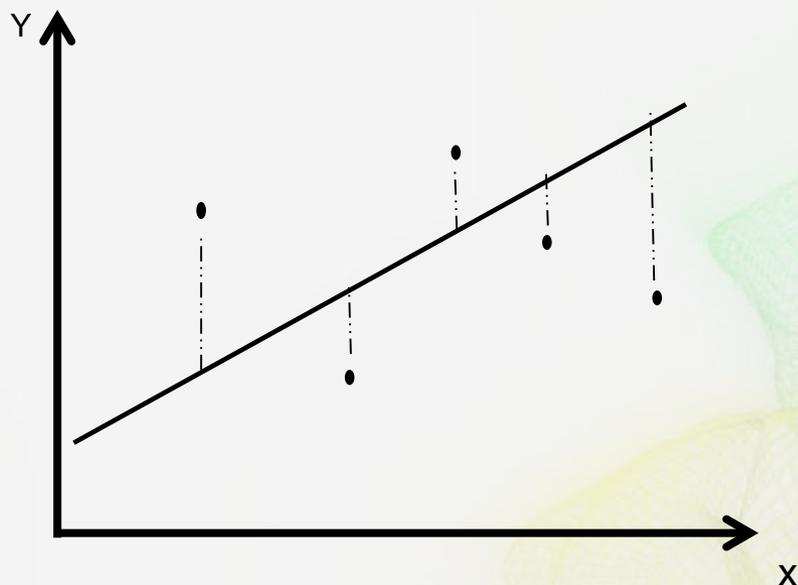
Fitting → Quanto bene il modello creato si adatta ai dati sperimentali usati per creare il modello stesso

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n$$



- PLS → Partial Least Square;

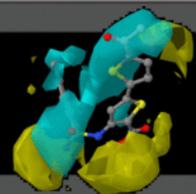
Tecnica basata sulla teoria dei minimi quadrati.



Fitting → Quanto bene il modello creato si adatta ai dati sperimentali usati per creare il modello stesso

Capacità predittiva → Quanto bene il modello creato spiega dati sperimentali non utilizzati per la creazione del modello

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n$$



COSTRUZIONE DI UN MODELLO

Scelta del training set

Allineamento delle molecole

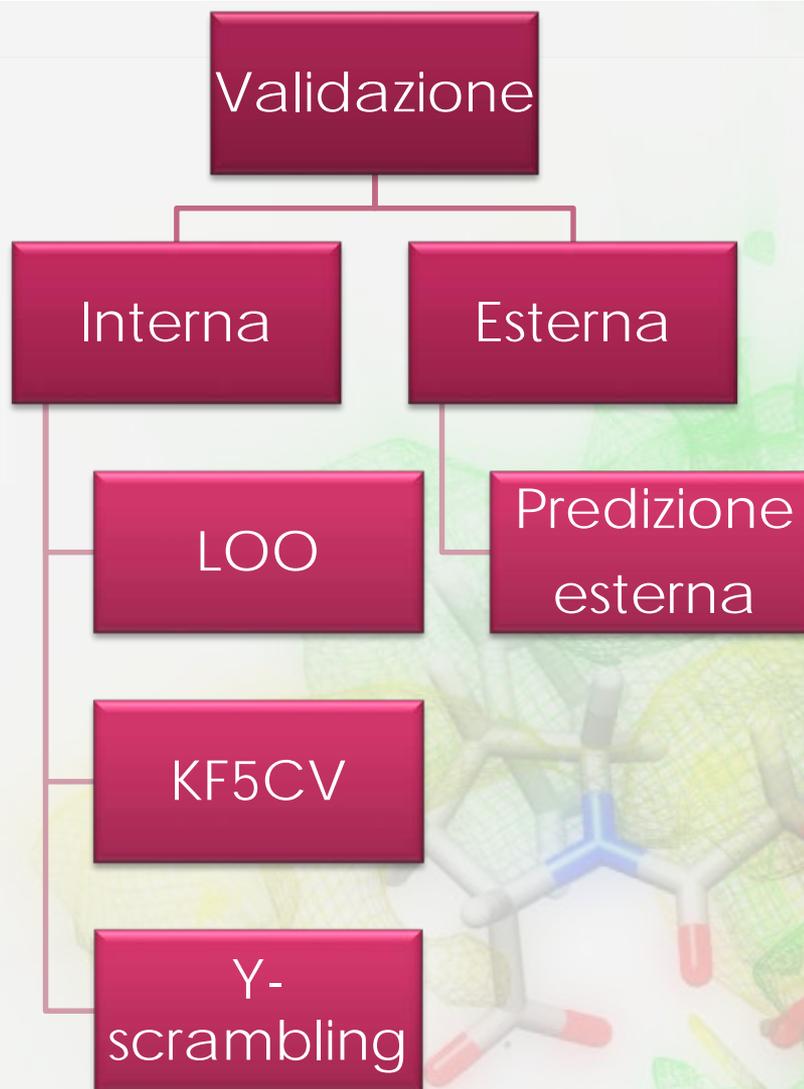
Calcolo dei campi di interazione molecolare (MIF)

Generazione del modello statistico

Validazione interna, esterna e y-scrambling



LA VALIDAZIONE



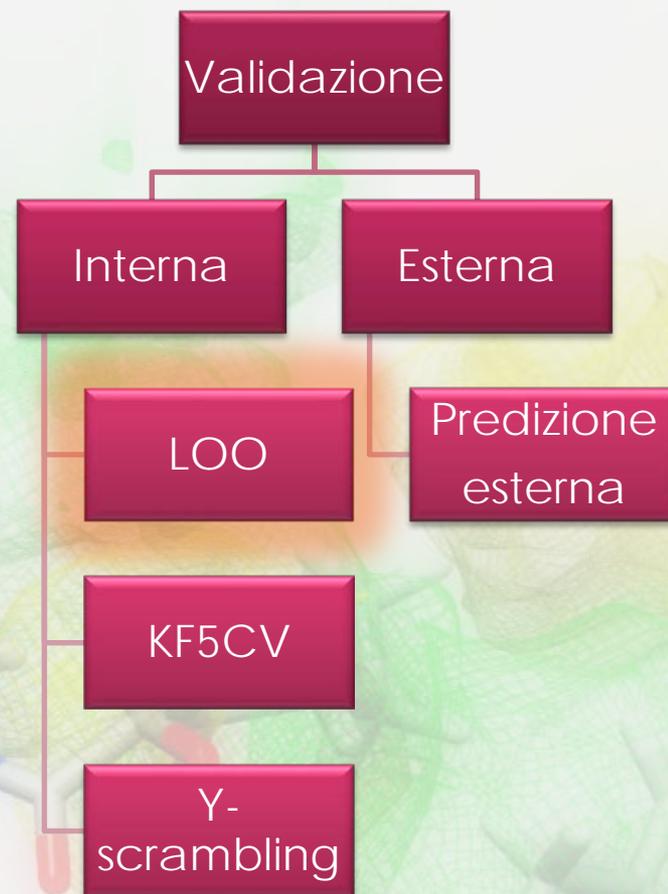


LA VALIDAZIONE

OBJECT 1
OBJECT 2
OBJECT 3
OBJECT 4
OBJECT 5
OBJECT 6
OBJECT 7
OBJECT 8
OBJECT 9
OBJECT 10
OBJECT 11
OBJECT 12
OBJECT 13
OBJECT 14
OBJECT 15

OBJECT 1
OBJECT 2
OBJECT 3
OBJECT 4
OBJECT 5
OBJECT 6
OBJECT 7
OBJECT 8
OBJECT 9
OBJECT 10
OBJECT 11
OBJECT 12
OBJECT 13
OBJECT 14
OBJECT 15

OBJECT 1
OBJECT 2
OBJECT 3
OBJECT 4
OBJECT 5
OBJECT 6
OBJECT 7
OBJECT 8
OBJECT 9
OBJECT 10
OBJECT 11
OBJECT 12
OBJECT 13
OBJECT 14
OBJECT 15



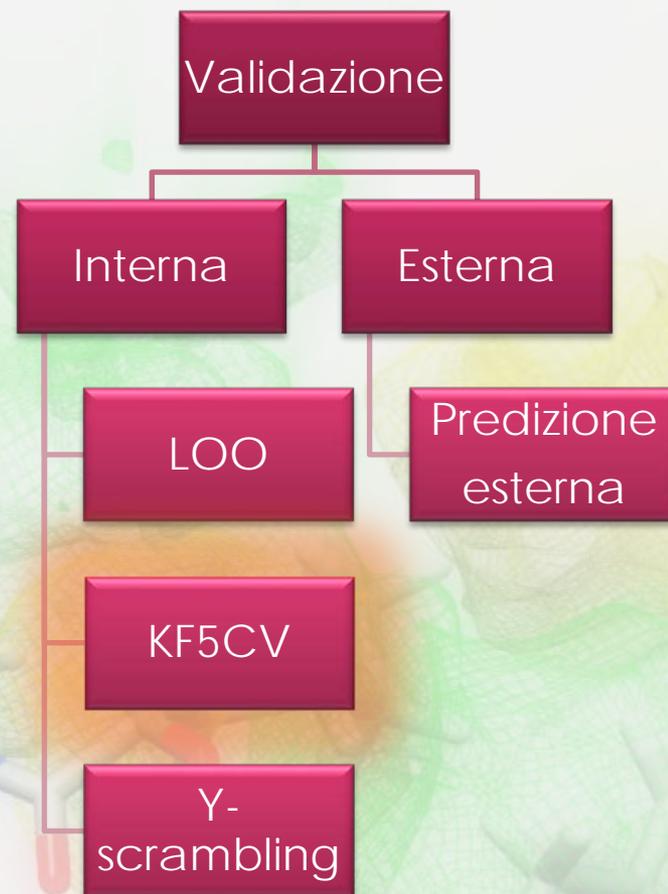


LA VALIDAZIONE

OBJECT 1
OBJECT 2
OBJECT 3
OBJECT 4
OBJECT 5
OBJECT 6
OBJECT 7
OBJECT 8
OBJECT 9
OBJECT 10
OBJECT 11
OBJECT 12
OBJECT 13
OBJECT 14
OBJECT 15

OBJECT 1
OBJECT 2
OBJECT 3
OBJECT 4
OBJECT 5
OBJECT 6
OBJECT 7
OBJECT 8
OBJECT 9
OBJECT 10
OBJECT 11
OBJECT 12
OBJECT 13
OBJECT 14
OBJECT 15

OBJECT 1
OBJECT 2
OBJECT 3
OBJECT 4
OBJECT 5
OBJECT 6
OBJECT 7
OBJECT 8
OBJECT 9
OBJECT 10
OBJECT 11
OBJECT 12
OBJECT 13
OBJECT 14
OBJECT 15





Validazione

Interna

Esterna

Y-Scrambling → "Shakera" le attività e calcola nuovamente il modello per n volte.

Se gli n modelli ottenuti sono paragonabili al modello che stiamo validando ci troviamo allora in un caso di change correlation, cioè il modello che stiamo validando è stato ottenuto per puro caso.



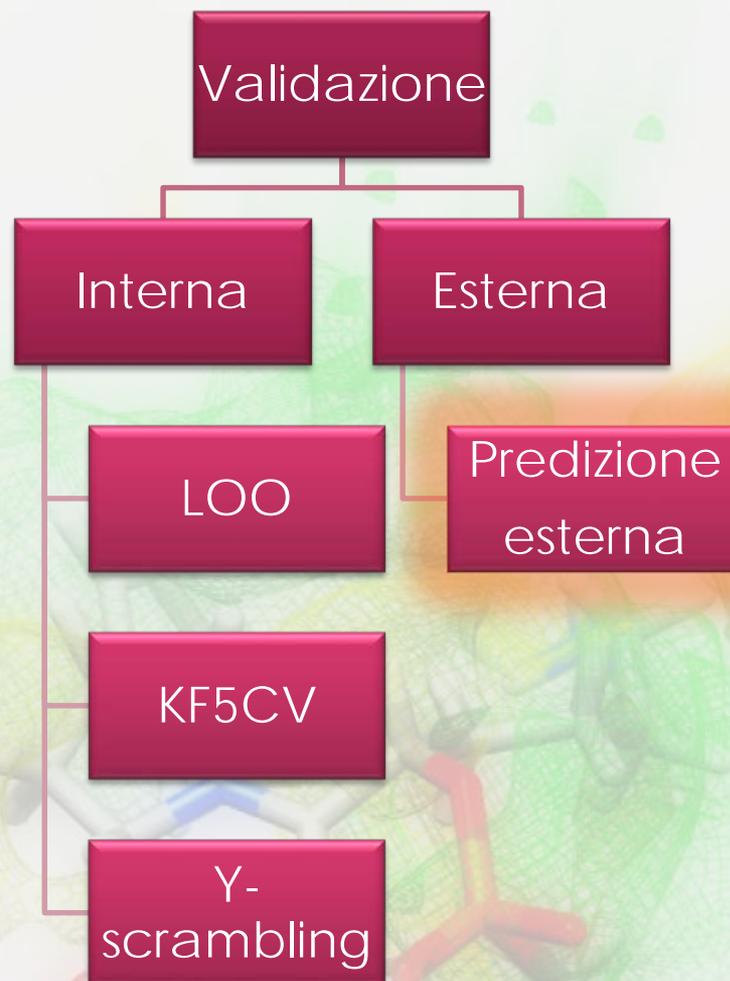


Validazione

Interna

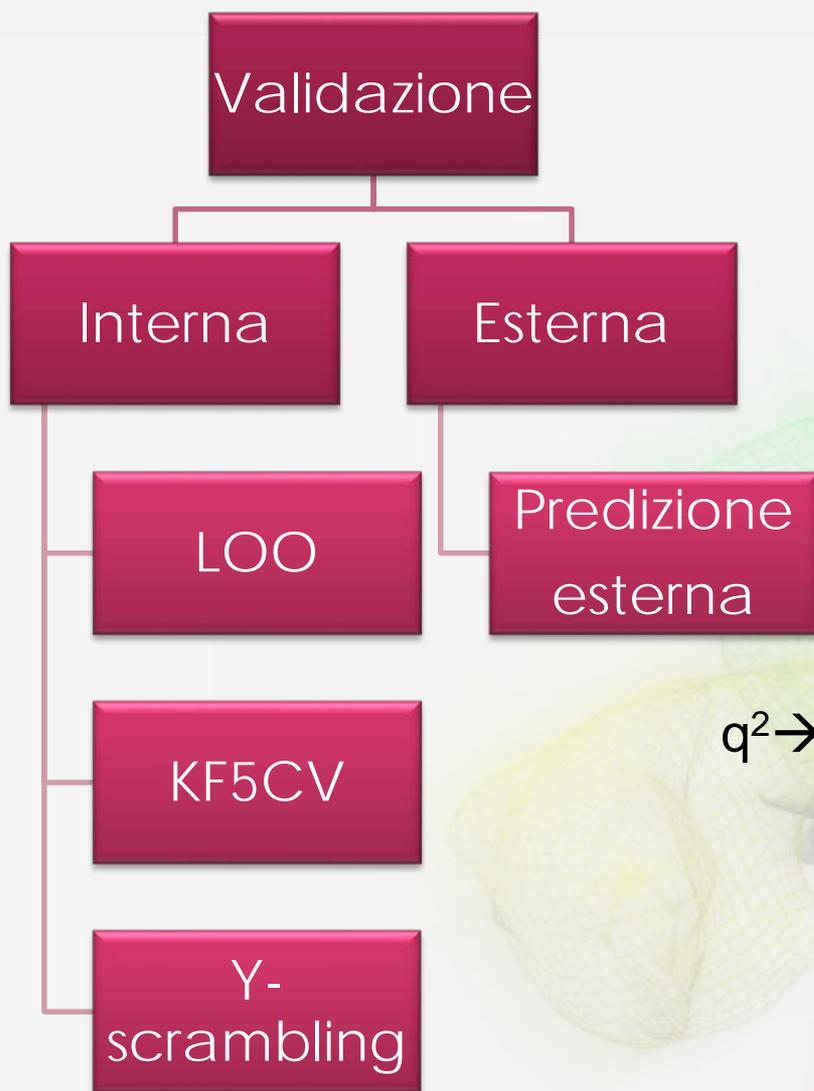
Esterna

Predizione Esterna → Si Divide il Data set in Training Set e il Test Set. Il Test set viene escluso durante la creazione del modello che viene invece usato poi per predirne le attività.





LA VALIDAZIONE

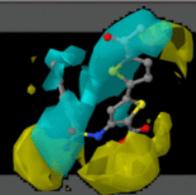


$$SDEP = \sqrt{\frac{\sum (y_{sperimentale} - y_{predetto})^2}{N}}$$

$$q^2 = 1 - \frac{\sum (y_{sperimentale} - y_{predetto})^2}{\sum (y_{sperimentale} - y_{medio})^2}$$

$q^2 \rightarrow$ quanto il modello è robusto
 $0 < q^2 < 1$

SDEP \rightarrow errore che il modello commette nella predizione



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



Sviluppo delle Contourn-Maps

Mesh areas



-



+

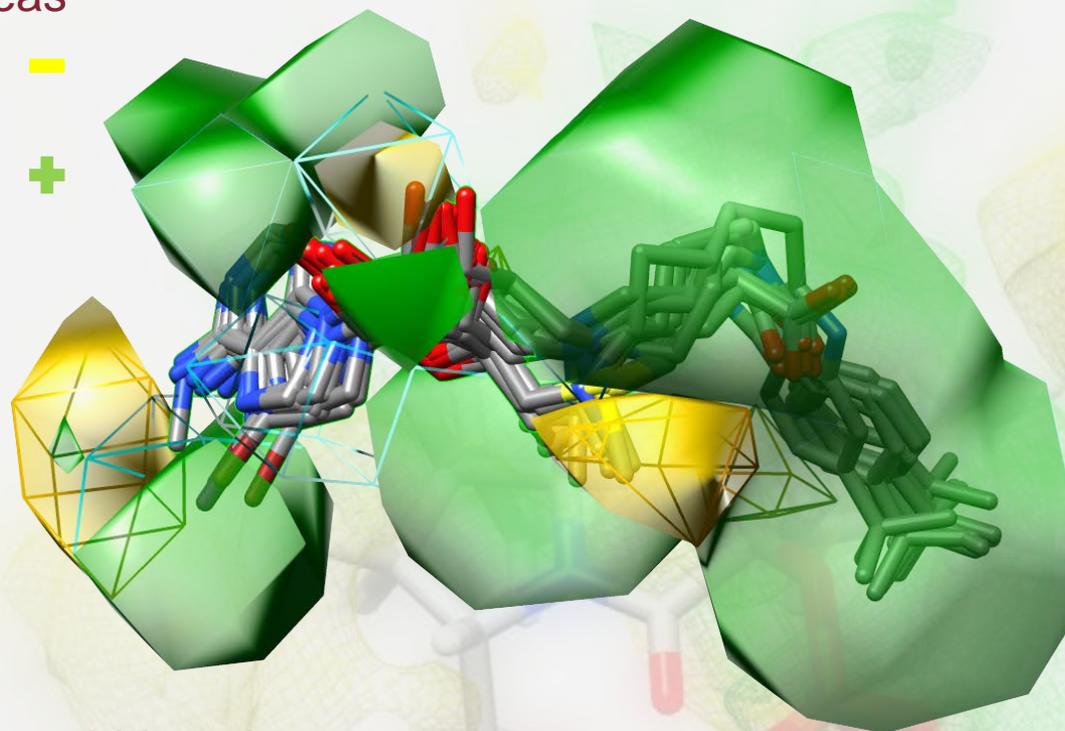
Solid areas



-



+



3-D QSAR: OA and N probes