

Heat Shock Protein 90 (HSP90): relazioni struttura-attività tridimensionali di tipo quantitativo (3-D QSAR) e valutazione della capacità predittiva.

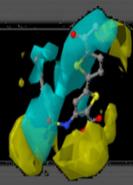


SAPIENZA
UNIVERSITÀ DI ROMA

Facoltà di Farmacia e Medicina
Corso di Laurea in Chimica e Tecnologia Farmaceutiche
Tesi Sperimentale in Chimica Farmaceutica
a.a. 2011/2012

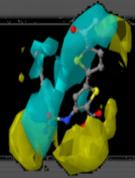
Laureando : Mario Esposito
Matricola: 316428

Relatore: Chiar.mo Prof. Rino Ragno



Heat Shock Protein 90

- Le HSP sono polipeptidi con funzione di chaperone molecolare ubiquitarie ad attività ATPasica
- Espressione aumentata in condizioni di stress cellulare e in diverse forme tumorali umane
- Interazione con centinaia di proteine *client* e co-chaperoni



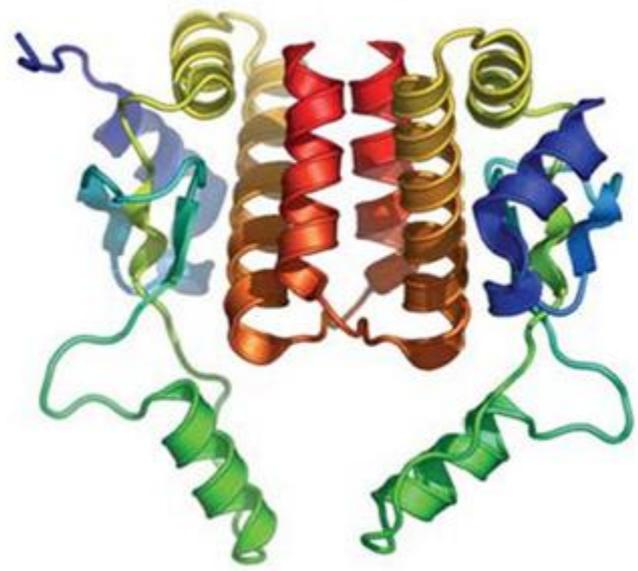
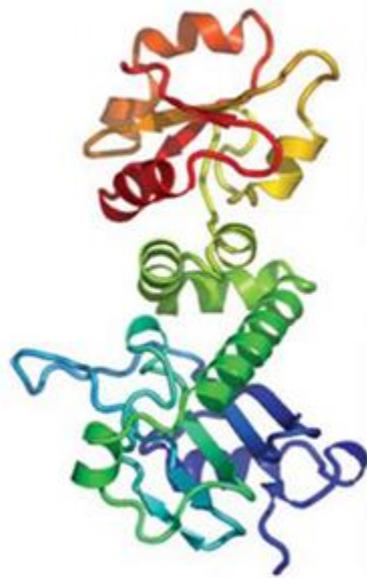
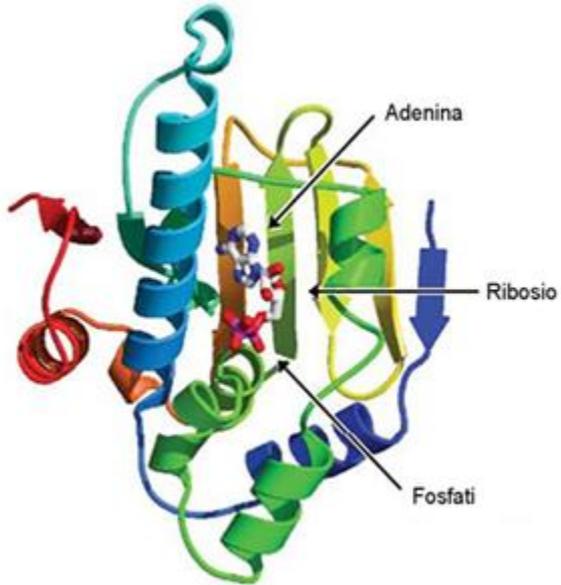
Heat Shock Protein 90

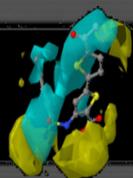


Dominio N-terminale
(25 Kda)

Dominio centrale
(35KDa)

Dominio C-terminale
(10KDa)





Progressione tumorale

1. Autosufficienza
signaling di
crescita

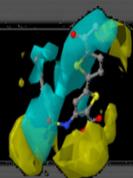
2. Insensibilità
signaling anti-
crescita

3. Evasione
apoptotica

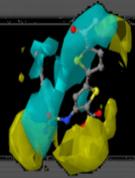
4. Replicazione
potenzialmente
illimitata

5. Angiogenesi
sostenuta

6. Invasione
tissutale e
metastasi



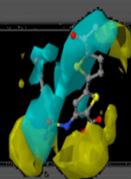
- **Tecniche Chimica Computazionale + Grafica Molecolare**
- **Studio 3-D QSAR su inibitori del dominio N-terminale di Hsp90: Costruzione e Validazione modello 3-D QSAR**
- **Interpretazione dei risultati**



Quantitative Structure-Activity Relationship (QSAR)

MIF
+
PLS

3-D QSAR



Comparative Molecular Field Analysis

J. Am. Chem. Soc. **1988**, *110*, 5959–5967

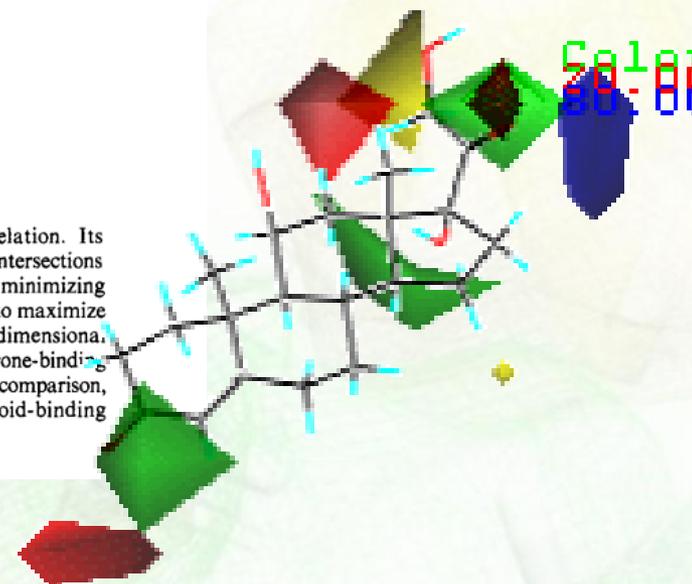
5959

Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins

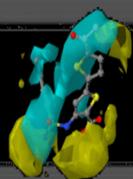
Richard D. Cramer, III,* David E. Patterson, and Jeffrey D. Bunce

Contribution from Tripos Associates, 1699 South Hanley Road, St. Louis, Missouri 63144. Received January 5, 1988

Abstract: Comparative molecular field analysis (CoMFA) is a promising new approach to structure/activity correlation. Its characteristic features are (1) representation of ligand molecules by their steric and electrostatic fields, sampled at the intersections of a three-dimensional lattice, (2) a new "field fit" technique, allowing optimal mutual alignment within a series, by minimizing the RMS field differences between molecules, (3) data analysis by partial least squares (PLS), using cross-validation to maximize the likelihood that the results have predictive validity, and (4) graphic representation of results, as contoured three-dimensional coefficient plots. CoMFA is exemplified by analyses of the affinities of 21 varied steroids to corticosteroid- and testosterone-binding globulins. Also described are the sensitivities of results to the nature of the field and the definition of the lattice and, for comparison, analyses of the same data using various combinations of other parameters. From these results, a set of ten steroid-binding affinity values unknown to us during the CoMFA analysis were well predicted.

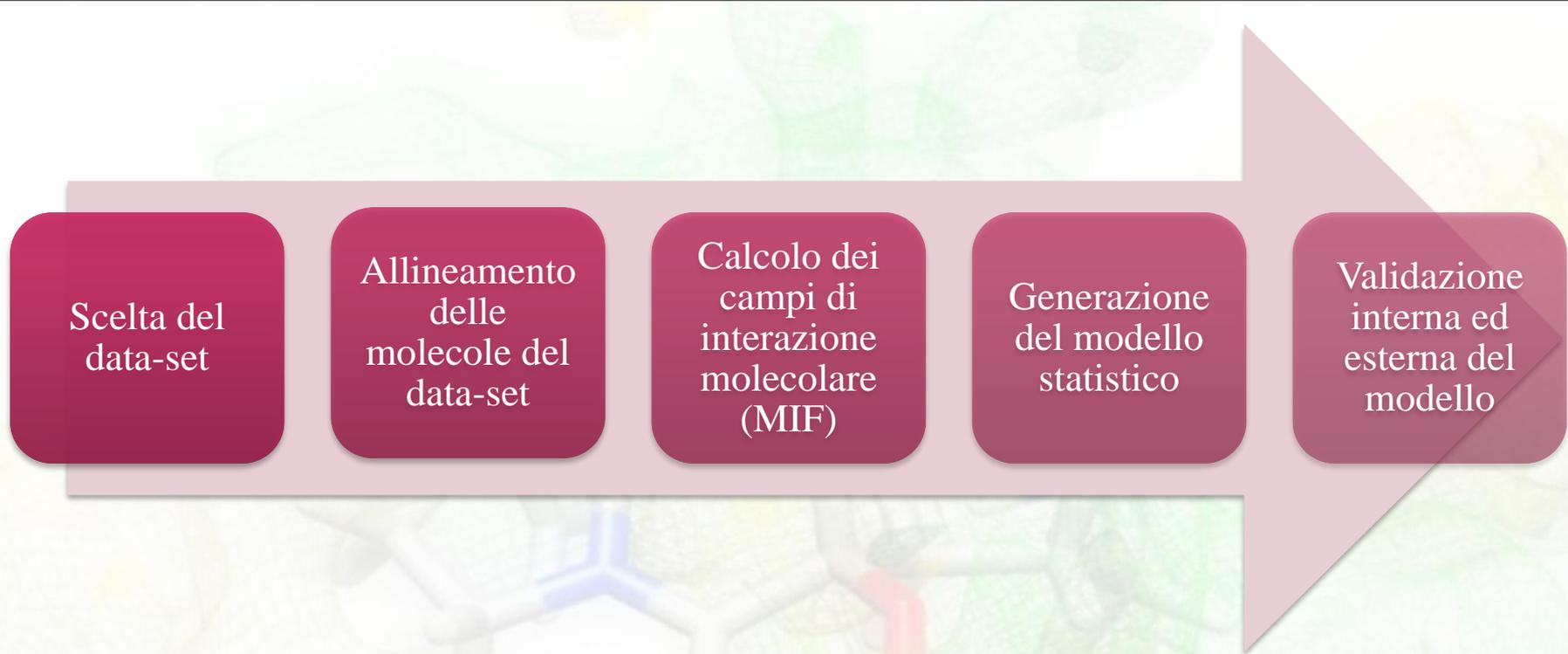


- ✓ Spaziatura nodi della griglia 2 Å
- ✓ Contributi sterici e elettrostatici
- ✓ Visualizzazione grafica mediante CoMFA Steric/Electrostatic contour maps



- **Spiegazione della correlazione tra proprietà chimico-fisiche dei composti e loro attività biologica**
- **Ottimizzazione molecole esistenti**
- **Predizione attività di composti non ancora sintetizzati**

Costruzione di un modello 3-D QSAR



Costruzione di un modello 3-D QSAR

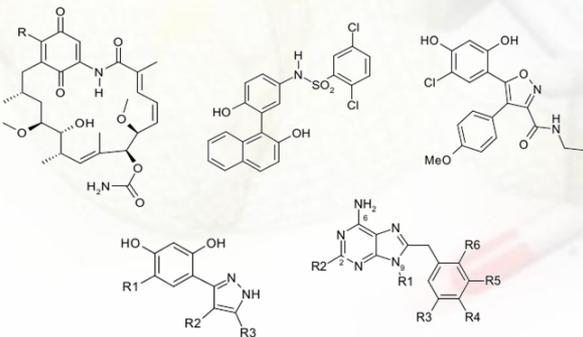
Scelta del data-set

Allineamento
delle molecole
del data-set

Calcolo dei
campi di
interazione
molecolare
(MIF)

Generazione
del modello
statistico

Validazione
interna ed
esterna del
modello



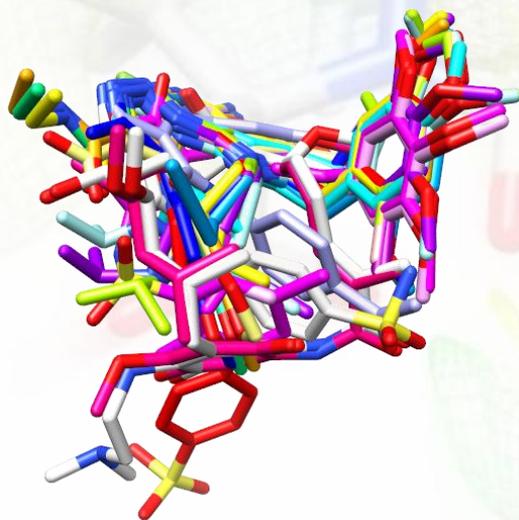
Scelta del
data-set

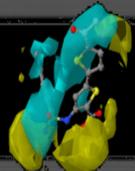
**Allineamento
delle molecole
del data-set**

Calcolo dei
campi di
interazione
molecolare
(MIF)

Generazione
del modello
statistico

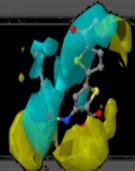
Validazione
interna ed
esterna del
modello



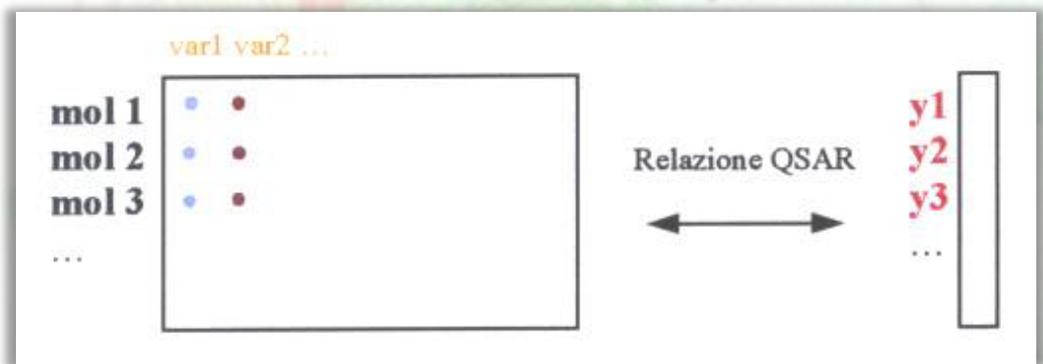
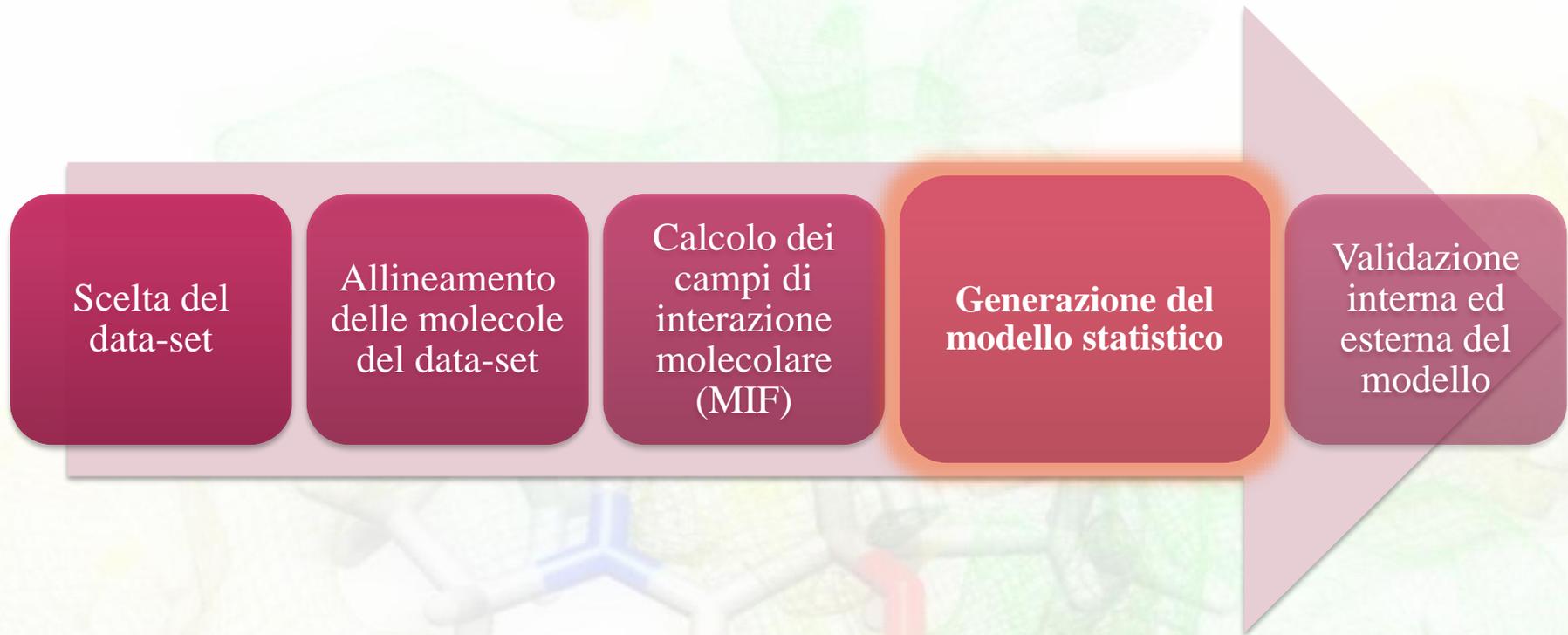


Costruzione di un modello 3-D QSAR





Costruzione di un modello 3-D QSAR



Scelta del
data-set

Allineamento
delle molecole
del data-set

Calcolo dei
campi di
interazione
molecolare
(MIF)

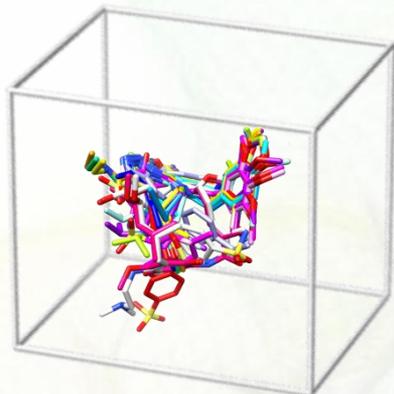
Generazione
del modello
statistico

Validazione
interna ed esterna
del modello

$$SDEP = \sqrt{\sum (y_{\text{exp}} - y_{\text{pred}})^2 / n - 1}$$
$$q^2 = 1 - \sum (y_{\text{exp}} - y_{\text{pred}})^2 / \sum (y_{\text{exp}} - y)^2$$



AutoGrid



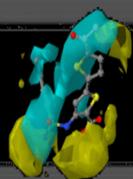
Calcolo dei Campi di Interazione Molecolare (MIF)

Probe	Significato
A	Carbonio aromatico
C	Carbonio alifatico
HD	Idrogeno che può formare legami H
NA	Azoto che può accettare legami H (amminico)
N	Azoto che non può accettare legami H (ammidico)
OA	Ossigeno che può accettare legami H
e	Carica elettrostatica unitaria
d	Desolvatazione

Analisi di regressione PLS (*Partial Least Square*)

- ✓ Modello grezzo
- ✓ Modello pretrattato
- ✓ Modello con selezione delle variabili GA
- ✓ Cross-validazione (LOO, K5F)
- ✓ Generazione mappe di contorno



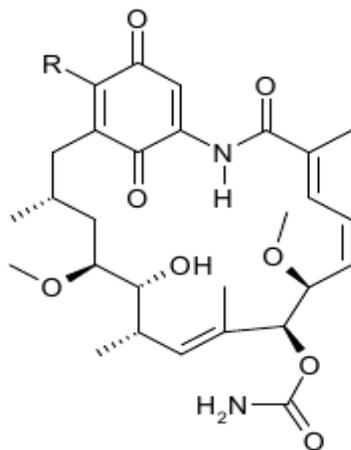


Training-set

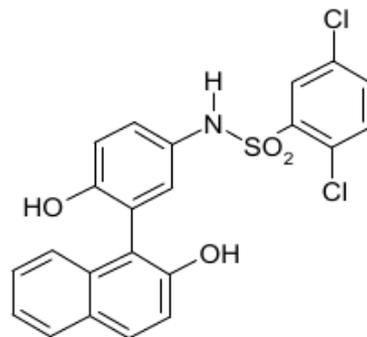
Codici PDB

1YET
1OSF
2BZ5
2UWD
1YC1
2BSM
2BYH
2BYI
2CCS
2CCT
2CCU
2CDD
1UY6
1UY7
1UY8
1UY9
1UYC
1UYD
1UYE
1UYF
1UYH
1UYI
1UYM
2H55
2FWY
2FWZ

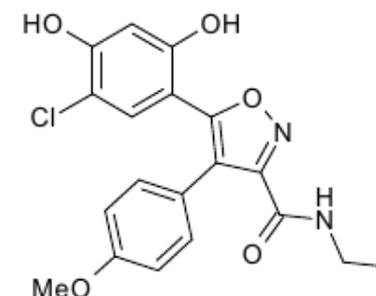
Ansamicine



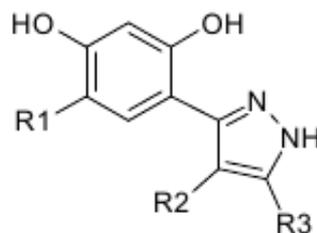
Benzenosulfonammidi



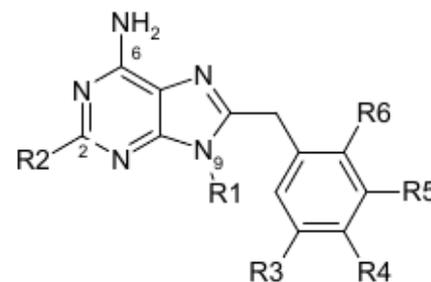
Isossazoli

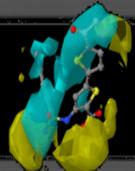


Pirazoli

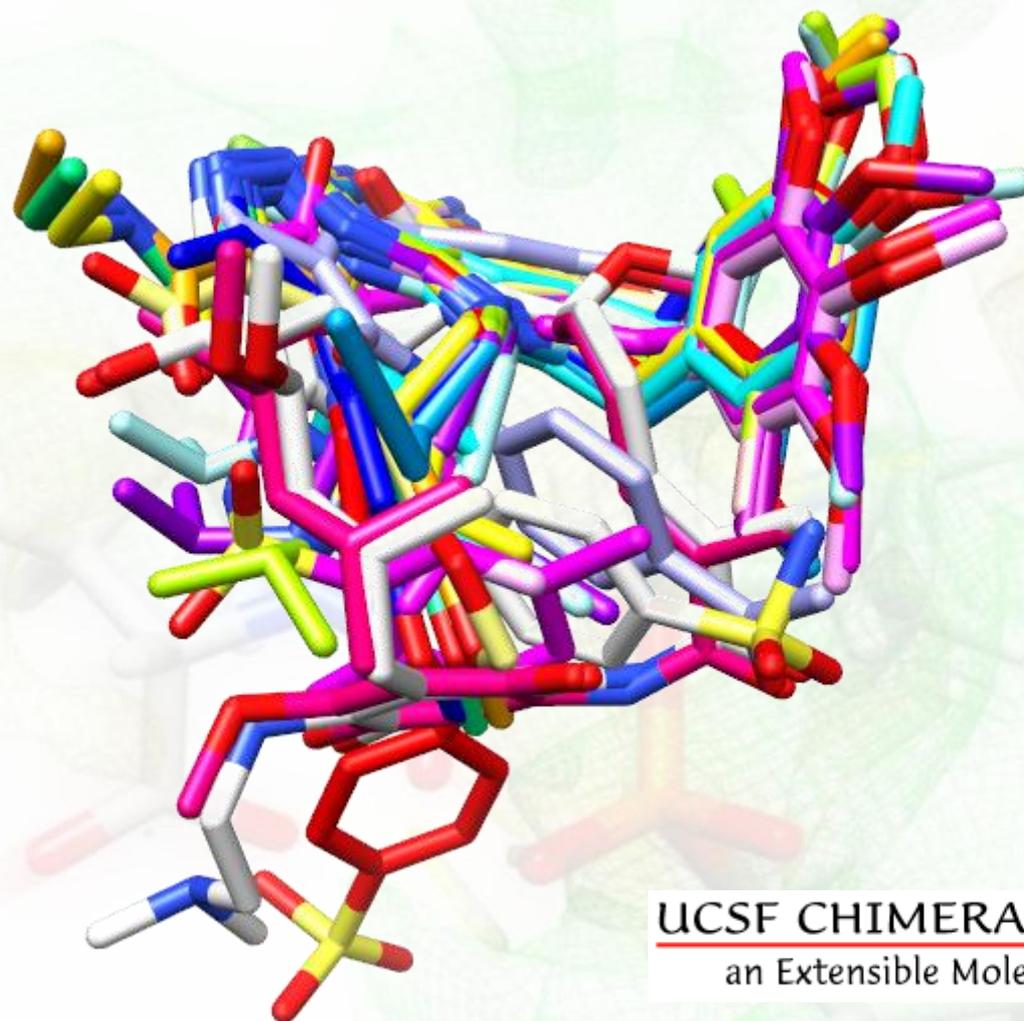


Purine

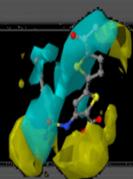




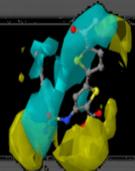
Allineamento del training-set



UCSF CHIMERA
an Extensible Molecular Modeling System



- **Metodi per validare il modello:**
- **Validazione interna (Cross-validazione)**
 - **Leave One Out CV**
 - **K5Fold CV**
- **Validazione esterna** → **Test-set esterno**
- **Conferma che il modello può essere usato per stimare l'attività biologica dei composti**



Risultati statistici dei modelli

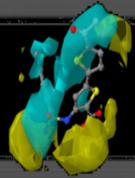
Probe	PC	r^2	q^2_{LOOCV}	q^2_{KF5CV}
A	3	0.91	0.69	0.63
C	3	0.91	0.69	0.62
HD	3	0.91	0.69	0.63
NA	3	0.90	0.68	0.61
N	3	0.90	0.68	0.62
OA	3	0.91	0.69	0.60
e	3	0.92	0.60	0.43
d	1	0.35	0.25	0.24

$$r^2 = 1 - \frac{\sum_{i=1}^N (Y_{\text{exp},i} - Y_{\text{calc},i})^2}{\sum_{i=1}^N (Y_{\text{exp},i} - \bar{Y})^2}$$

$$0 \leq r^2 \leq 1$$

$$q^2 = 1 - \frac{\sum_{i=1}^N (Y_{\text{exp},i} - Y_{\text{pred},i})^2}{\sum_{i=1}^N (Y_{\text{exp},i} - \bar{Y})^2}$$

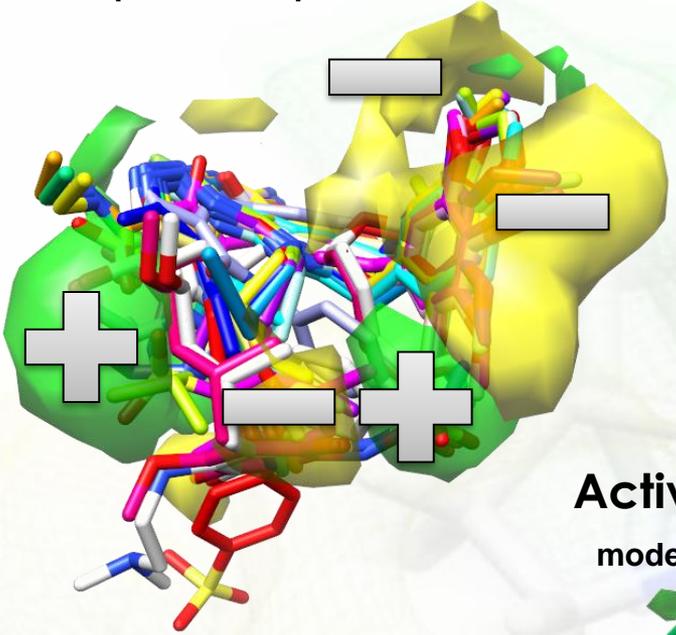
$$0 \leq q^2 \leq 1$$



Risultati grafici: mappe di contorno

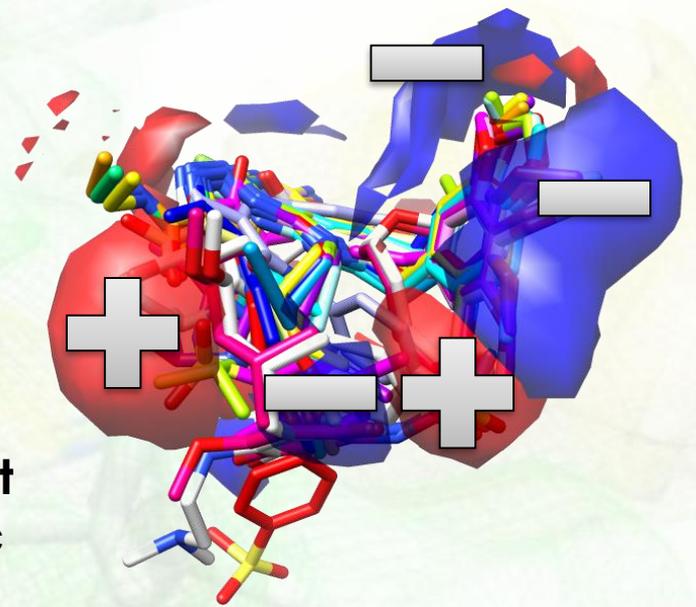
CoMFA like map

modello pretrattato, *probe "A"* 3PC "Contributo sterico"



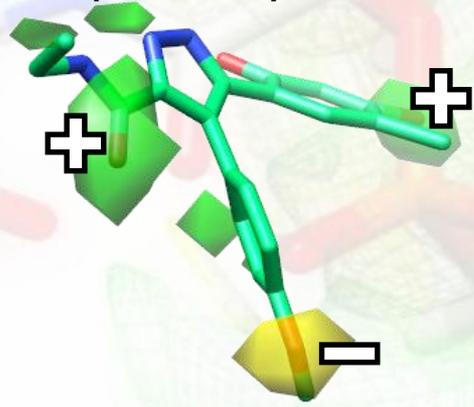
PLS Coefficients Plot

modello pretrattato, *probe "HD"* 3PC

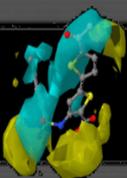


Activity Contribution Plot

modello pretrattato, *probe "A"* 3PC

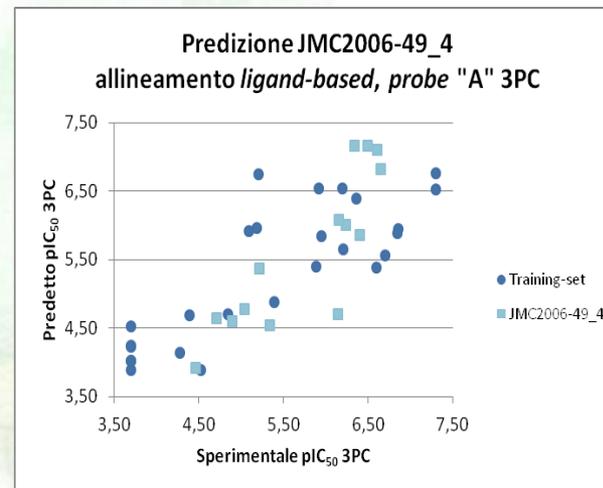
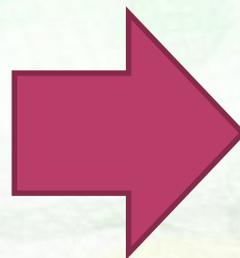
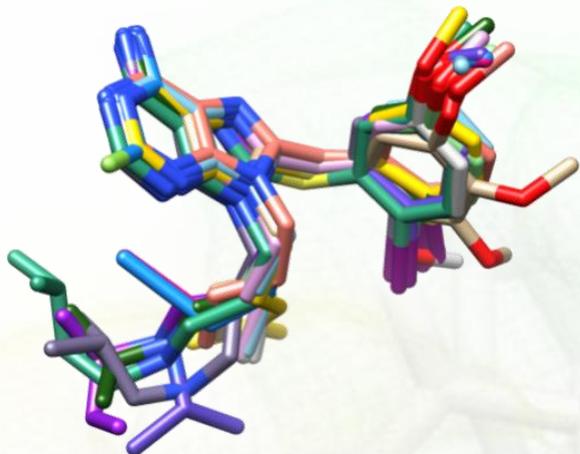


Una delle molecole più attive 2BSM ($pIC_{50} = 6,85$)



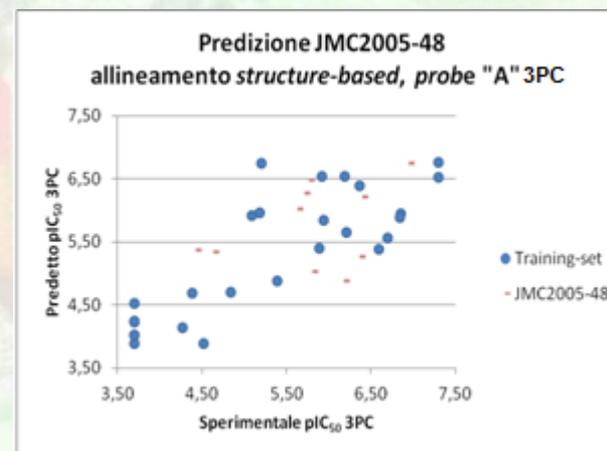
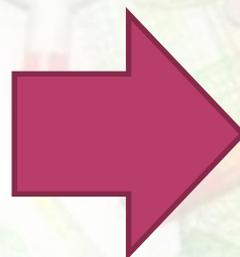
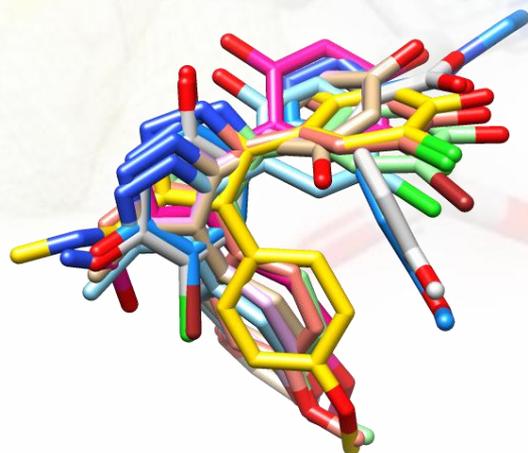
Allineamento del test-set esterno

Allineamento *ligand-based* di JMC2006-49_4 (software Surflex-Sim)

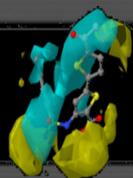


SDEP = 0,59

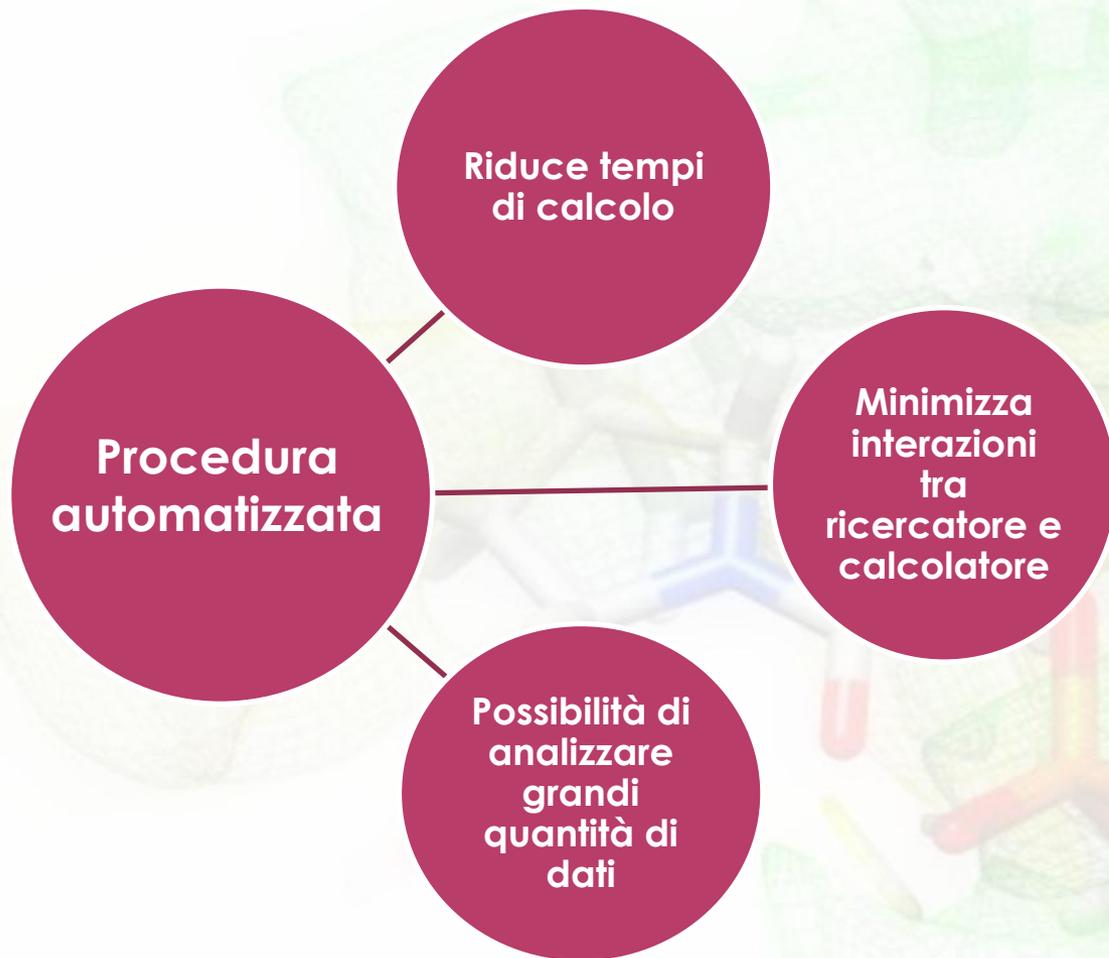
Allineamento *structure-based* di JMC2005-48 (software AutoDock)



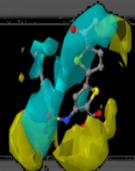
SDEP = 0,77



Tecnica 3-D QSAR con il metodo 3-D QSAutogrid/R



- Efficacia e robustezza del metodo 3-D QSAutogrid/R
- Costruzione modelli e predizione molecole esterne
- Buoni risultati ottenuti -> Creazione di un database online di modelli 3-D QSAR per predire l'attività biologica di molecole incognite



Conclusioni

www.3d-qsar.com

3D QSAR www.3d-qsar.com
Virtual Screening, Quantitative Structure-Activity Relationship, Business Event

rcsd www.rcsd.it
www.rcsd.it

Home Launch Predictor **Datasets** About Change Log

You are inspecting model **STERIODS_e**

Model Description

The STERIODS dataset was taken from J. Mol. Model. 1992, 4, 276 - 283. Paper abstract: These sets of molecules have been used to study the conventional CoMFA procedure. For all the three test sets, the resulting q^2 values were observed to vary simply because of the change in the orientation or placement of the aligned molecules. The reason is believed to rest in the imperfect sampling of the molecular field. We have introduced two new strategies, all-orientation search (AOS) and all-placement search (APS), to optimize the sampling process. By rotating and translating the molecule aggregates within the grid systematically, all the possible samplings of the molecular field are tested and subsequently the one with the highest q^2 value can be picked out. We have also demonstrated that the combined application of AOS/APS with OCLIFE procedure can yield results better than the ones by using them respectively.

STERIODS Models

- STERIODS_e
- STERIODS_c
- STERIODS_a
- STERIODS_b
- STERIODS_i
- STERIODS_eq

Other Datasets

- AC
- ACE
- ACE_e
- N
- AD
- AT

Training Set

Show	Mol id	Exp. [pIC]
●	1370	9.740
○	1379	9.204
○	1380	9.170
○	1047	9.114
○	1156	8.833
○	1372	8.176
○	1386	7.819
○	1381	7.462

Maps Table

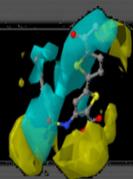
%	0	10	20	30	40	50	60	70	80	90	100
PLS_Coeff	●	○	○	○	○	○	○	○	○	○	○
CoMFA_Map	●	○	○	○	○	○	○	○	○	○	○

Interpret a 3-D QSAR model requires to look to different grid plots, in which the chemometric information obtained by the model is represented in the real space around the molecules. In the active plot window the training set molecules can be loaded and arranged from the grid plot can be displayed. In the actual version of the server you can have two important plots: PLS_Coeff and CoMFA like plots.

PLS Coefficients: The values represented here summarize the property weighted contributions of each PC, up to a certain level, however in this plot the signs of the coefficients can induce to errors, coefficients have opposite meaning depending on the fact that the compound produces positive or negative field values in this area. Cyan maps are the positive coefficients and yellow are the negatives.

CoMFA like maps: The values represented here are the PLS coefficients * standard deviation products and are somehow the same plots obtained in a standard CoMFA. The green maps (positives) give areas where an increase of steric hindrance is directly correlated with the biological activity, while orange maps (negatives) are inversely correlated with the steric hindrance.

© 2011 3D QSAR SERVER provided by R.C.S.D.



Ringraziamenti

**Ringrazio il
Prof. Rino Ragno e tutto il
gruppo di ricerca dell'RCMD
per il supporto ricevuto durante
il corso della mia tesi di laurea**

Grazie per l'attenzione