Procedura computazionale per la generazione automatica di modelli 3-D QSAR con approccio quasi-sistematico

SAPIENZA
UNIVERSITÀ DI ROMA

Facoltà di Farmacia e Medicina
Corso di Laurea in Farmacia
Tesi Sperimentale in Chimica Farmaceutica
a.a. 2014/2015

Laureanda: Sonia Ginersio
Matricola: 1382618

Relatore: prof. Rino Ragno

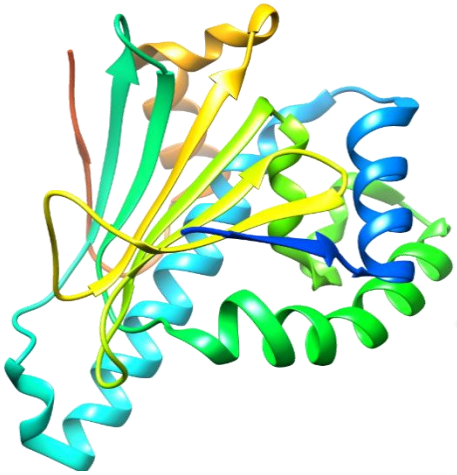
Introduzione

by www.RCMD.it

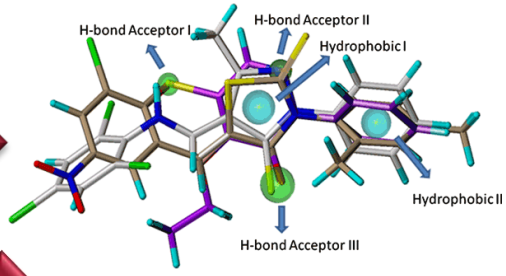
DRUG DESIGN

DISPONIBILITA' STRUTTURA 3D

SÌ



NO



STRUCTURE-BASED

- DOCKING
- DINAMICA MOLECOLARE

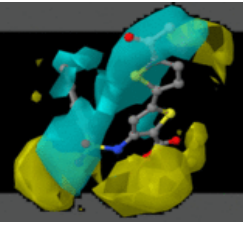
CONFORMAZIONI

LIGAND-BASED

- 3-D QSAR
- SIMILARITÀ MORFOLOGICA

ALLINEAMENTO

GENERAZIONE DI MODELLI STATISTICO-MATEMATICI VALIDI



Protocollo

by www.RCMD.it

PROTOCOLLO VALIDAZIONE

Allineamento dei ligandi cristallizzati mediante la metodologia structure-based

Chimera-MatchMaker

Generazione dei MIF

Creazione del modello 3-D QSAR

Validazione del modello

LOO | **Yscrambling**

Allineamento dei ligandi cristallizzati ligand-based

Surflex | **Kcombu** | **Shaep**

Open3Dalign | **Align-it**

Shape-it

Generazione dei MIF

Creazione del modello 3-D QSAR

Validazione del modello

LOO | **Yscrambling**

Analisi e confronto dei dati numerici statistici ottenuti

OGGETTO DI STUDIO:
Hsp90, VEGFR2

Protocollo

by www.RCMD.it

PROTOCOLLO VALIDAZIONE

Allineamento dei ligandi cristallizzati mediante la metodologia structure-based

Chimera-MatchMaker

Generazione dei MIF

Creazione del modello 3-D QSAR

Validazione del modello

LOO | **Yscrambling**

Allineamento dei ligandi cristallizzati ligand-based

Surflex | **Kcombu** | **Shaep**

Open3Dalign | **Align-it**

Shape-it

Generazione dei MIF

Creazione del modello 3-D QSAR

Validazione del modello

LOO | **Yscrambling**

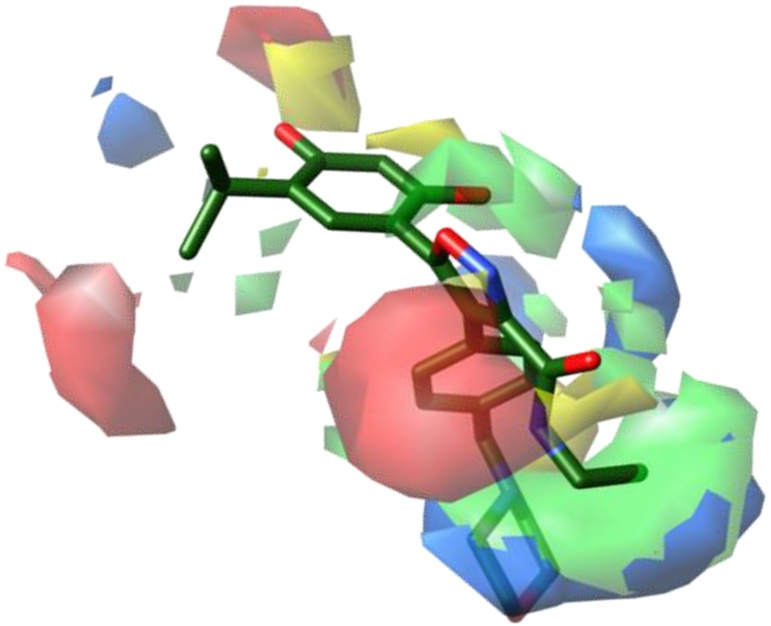
Analisi e confronto dei dati numerici statistici ottenuti

OGGETTO DI STUDIO:
Hsp90, VEGFR2

Costruzione di un modello 3-D QSAR

by www.RCMD.it

- SCELTA DEL TRAINING SET
- ALLINEAMENTO DELLE MOLECOLE
- CALCOLO DEL CAMPI DI INTERAZIONE MOLECOLARE (MIF)
- GENERAZIONE DEL MODELLO STATISTICO
- VALIDAZIONE



Costruzione di un modello 3-D QSAR

by www.RCMD.it

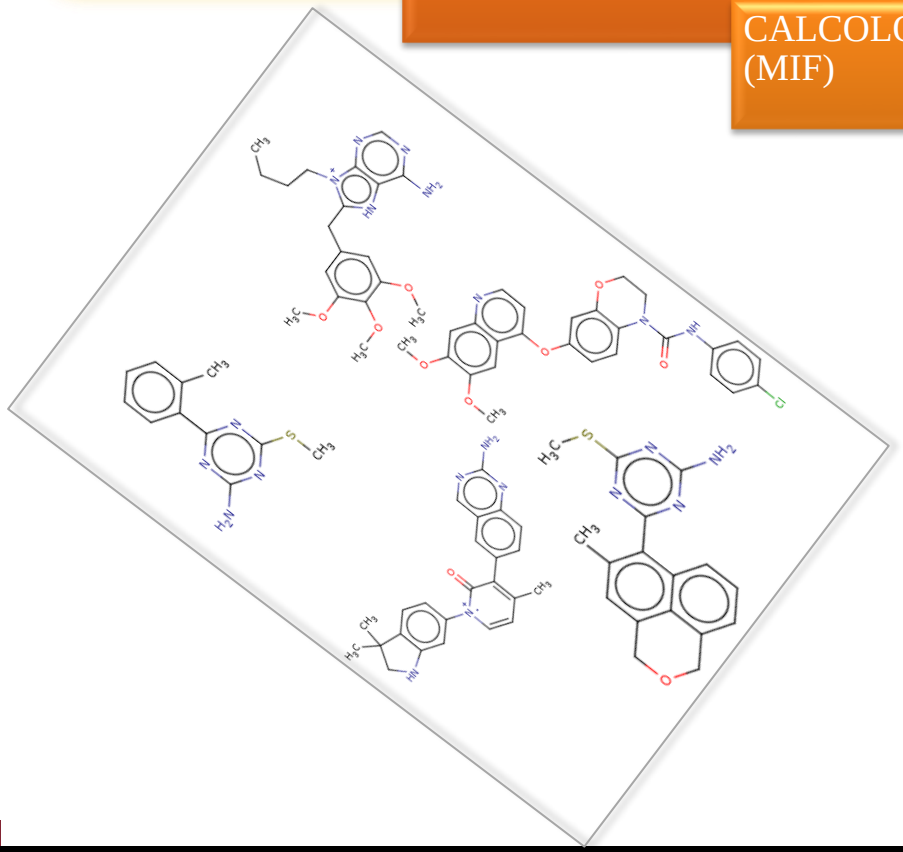
SCELTA DEL TRAINING SET

ALLINEAMENTO DELLE MOLECOLE

CALCOLO DEL CAMPI DI INTERAZIONE MOLECOLARE (MIF)

GENERAZIONE DEL MODELLO STATISTICO

VALIDAZIONE



Costruzione di un modello 3-D QSAR

by www.RCMD.it

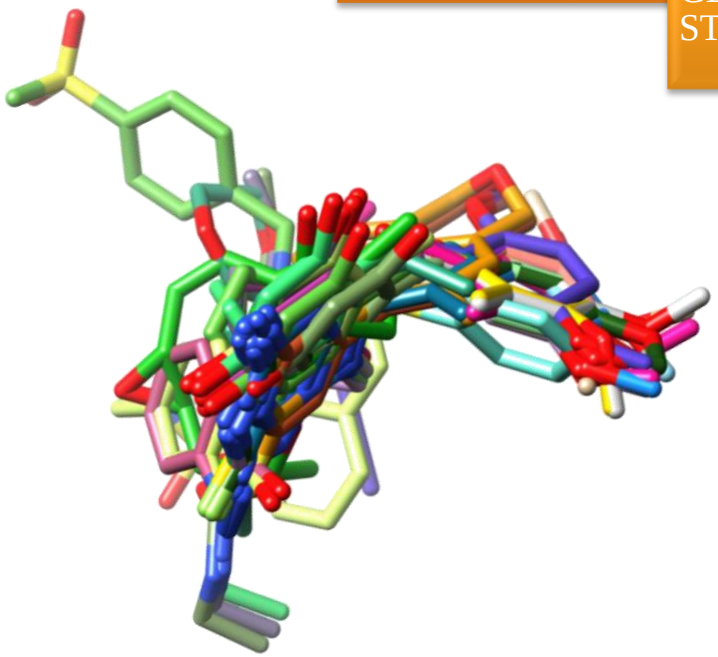
SCELTA DEL TRAINING SET

ALLINEAMENTO DELLE MOLECOLE

CALCOLO DEL CAMPI DI INTERAZIONE MOLECOLARE (MIF)

GENERAZIONE DEL MODELLO STATISTICO

VALIDAZIONE



Costruzione di un modello 3-D QSAR

by www.RCMD.it

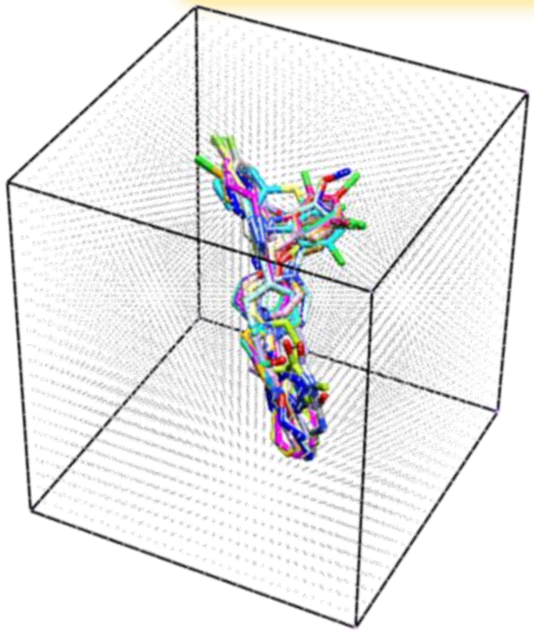
SCELTA DEL TRAINING SET

ALLINEAMENTO DELLE MOLECOLE

CALCOLO DEL CAMPI DI INTERAZIONE MOLECOLARE (MIF)

GENERAZIONE DEL MODELLO STATISTICO

VALIDAZIONE



Costruzione di un modello 3-D QSAR

by www.RCMD.it

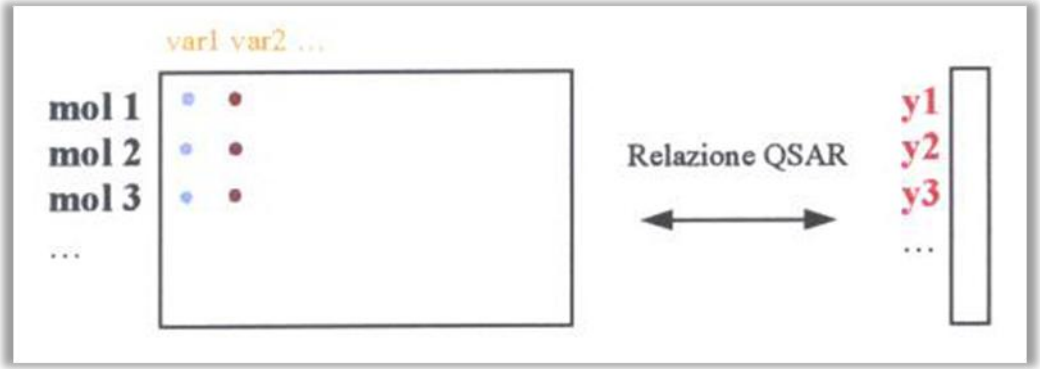
SCELTA DEL TRAINING SET

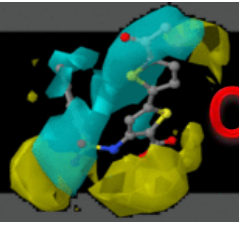
ALLINEAMENTO DELLE MOLECOLE

CALCOLO DEL CAMPI DI INTERAZIONE MOLECOLARE (MIF)

GENERAZIONE DEL MODELLO STATISTICO

VALIDAZIONE





Costruzione di un modello 3-D QSAR

by www.RCMD.it

- SCELTA DEL TRAINING SET
- ALLINEAMENTO DELLE MOLECOLE
- CALCOLO DEL CAMPI DI INTERAZIONE MOLECOLARE (MIF)
- GENERAZIONE DEL MODELLO STATISTICO
- **VALIDAZIONE**

$$SDEP = \frac{\sqrt{\sum(y_{exp} - y_{pred})^2}}{n} - 1$$

$$q^2 = 1 - \sum(y_{exp} - y_{pred})^2 \Big/ \sum\left(y_{exp} - (y)\right)^2$$

Relazioni quantitative struttura-attività

(QSAR)

MIF
+
PLS

3-D QSAR

Ricerca farmacoforica
Ottimizzazione delle molecole esistenti
Predizione delle attività di nuove entità chimiche
Progettazione di nuove molecole

Protocollo

by www.RCMD.it

PROTOCOLLO VALIDAZIONE

Allineamento dei ligandi cristallizzati mediante la metodologia structure-based

Chimera-MatchMaker

Generazione dei MIF

Creazione del modello 3-D QSAR

Validazione del modello

LOO | Yscrambling

Allineamento dei ligandi cristallizzati ligand-based

Surflex | Kcombu | Shaep | Open3Dalign | Align-it | Shape-it

Generazione dei MIF

Creazione del modello 3-D QSAR

Validazione del modello

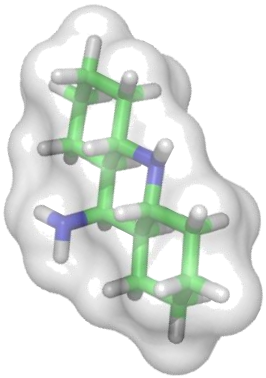
LOO | Yscrambling

Analisi e confronto dei dati numerici statistici ottenuti

OGGETTO DI STUDIO:
Hsp90, VEGFR2

ALGORITMI DI ALLINEAMENTO

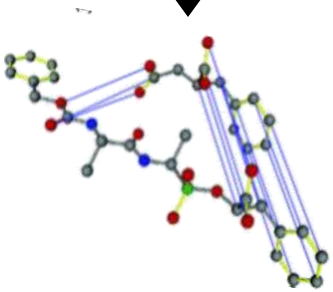
OTTIMIZZAZIONE DEI VOLUMI SOVRAPPPOSTI



ALIGN-IT

SHAPE-IT

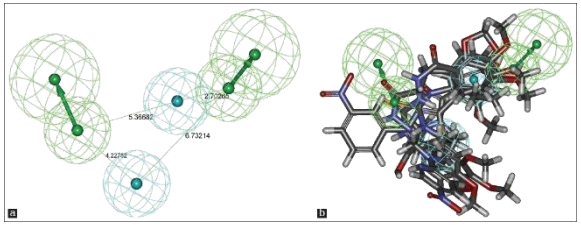
CONNETTIVITÀ DEI GRAFI



KCOMBU

OPEN 3D ALIGN

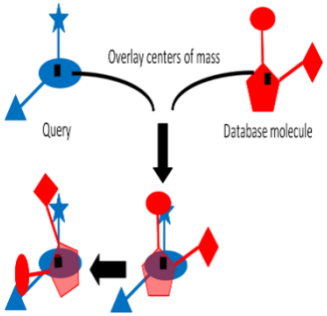
BASATO SUL FARMACOFORO



OPEN 3D ALIGN

ALIGN IT

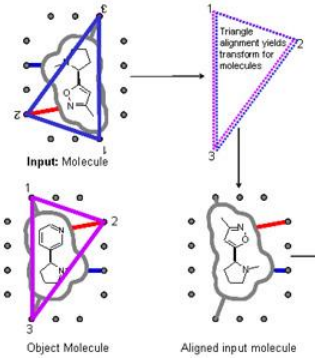
SHAPE



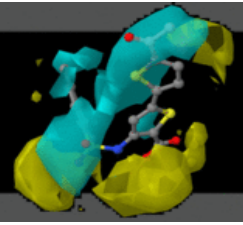
SHAEP

SHAPE IT

SIMILARITÀ MORFOLOGICA



SURFLEX



Procedura

by www.RCMD.it

PROTEIN DATA BANK

⇩

COMPLESSI LIGANDO-PROTEINA

⇩ **MINIMIZZAZIONE**
ESTRAZIONE LIGANDI

Data set sperimentale

TRAINING SET

N° composti

Allineamento SB

MatchMaker

⇩ **MIF**

- **MODELLI 3D-QSAR DI RIFERIMENTO**
- **VISUALIZZAZIONE E INTERPRETAZIONE DELLE MAPPE**

q^2

Best q^2

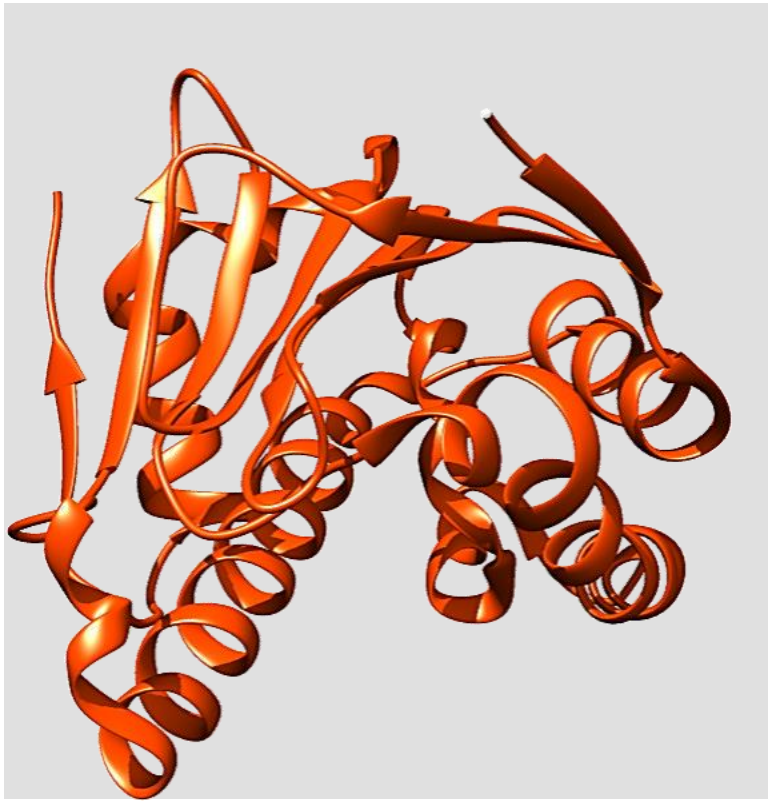
Allineamento LB

QSAlign.py

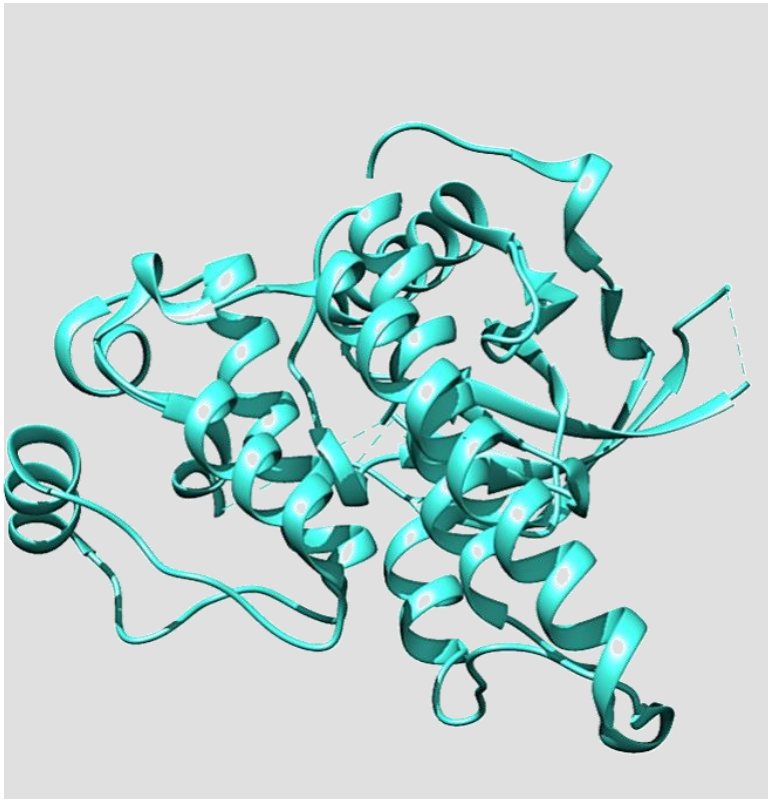
MIF ⇩

- **N° allineamenti x N° probe = N° MODELLI 3D-QSAR**
- **VISUALIZZAZIONE E INTERPRETAZIONE DELLE MAPPE**

Dataset

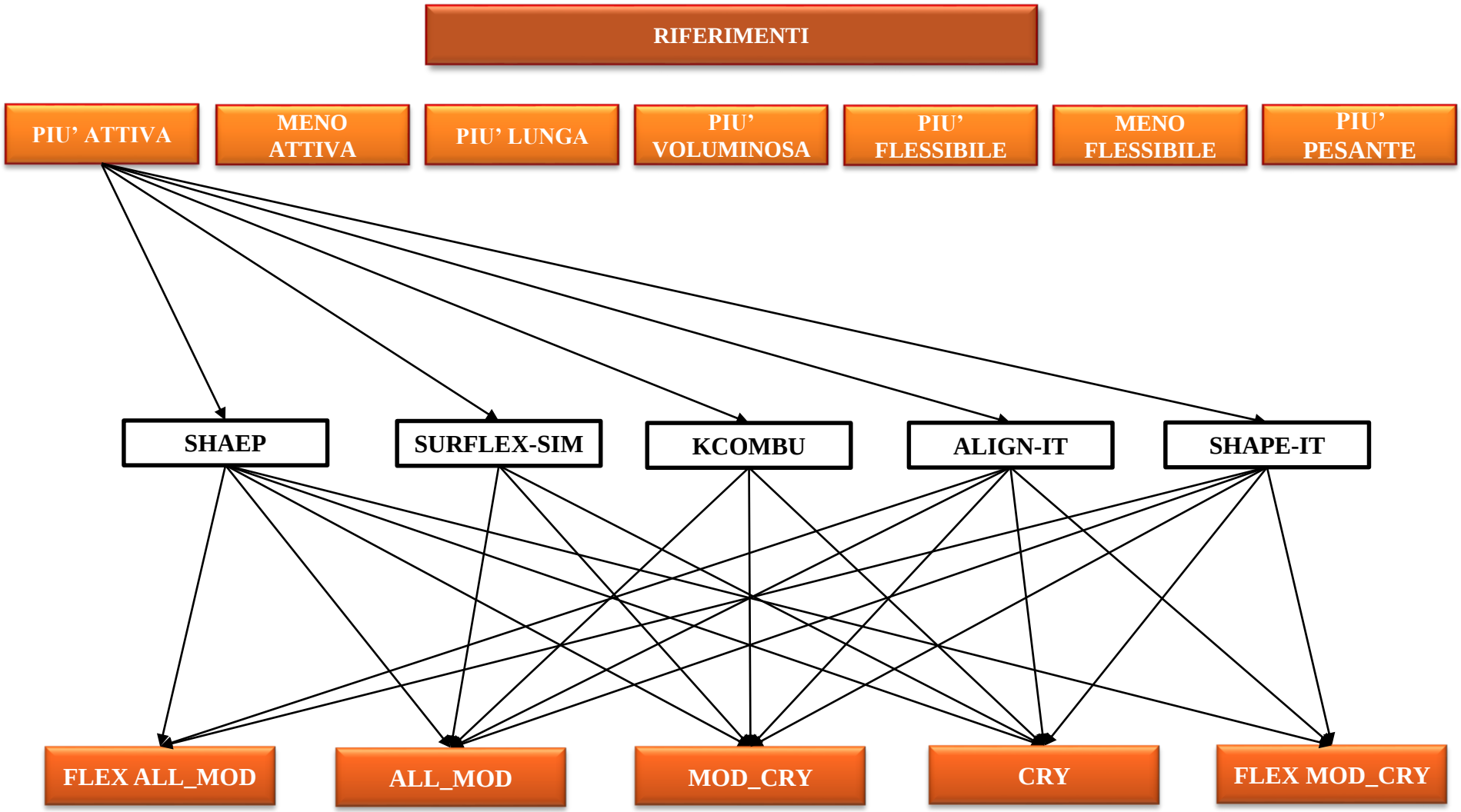


Hsp90
Heat Shock Protein
(PDB CODE:1UY7)



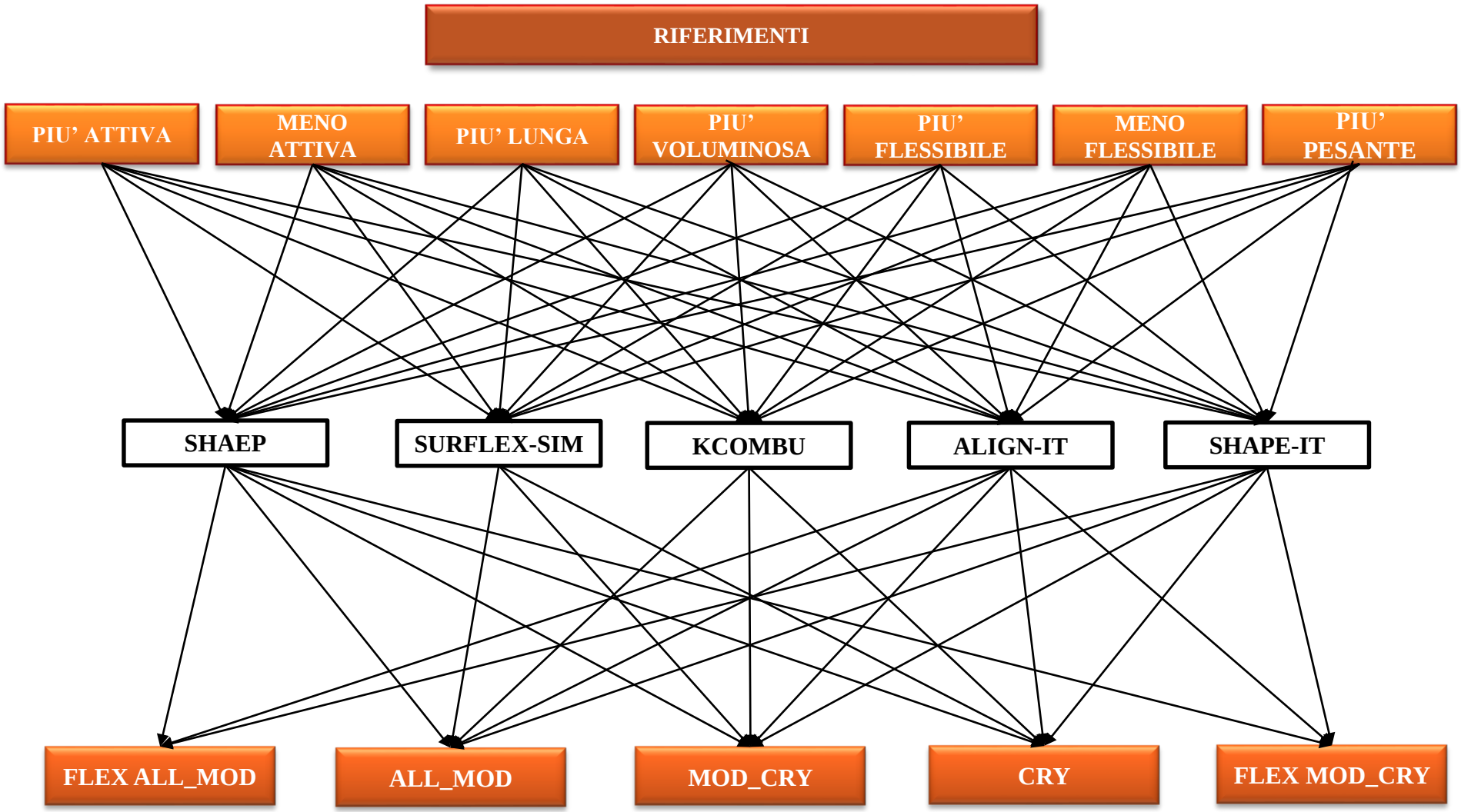
VEGFR2
Vascular Endothelial Growth Factor Receptor
(PDB CODE:3DTW)

Allineamento



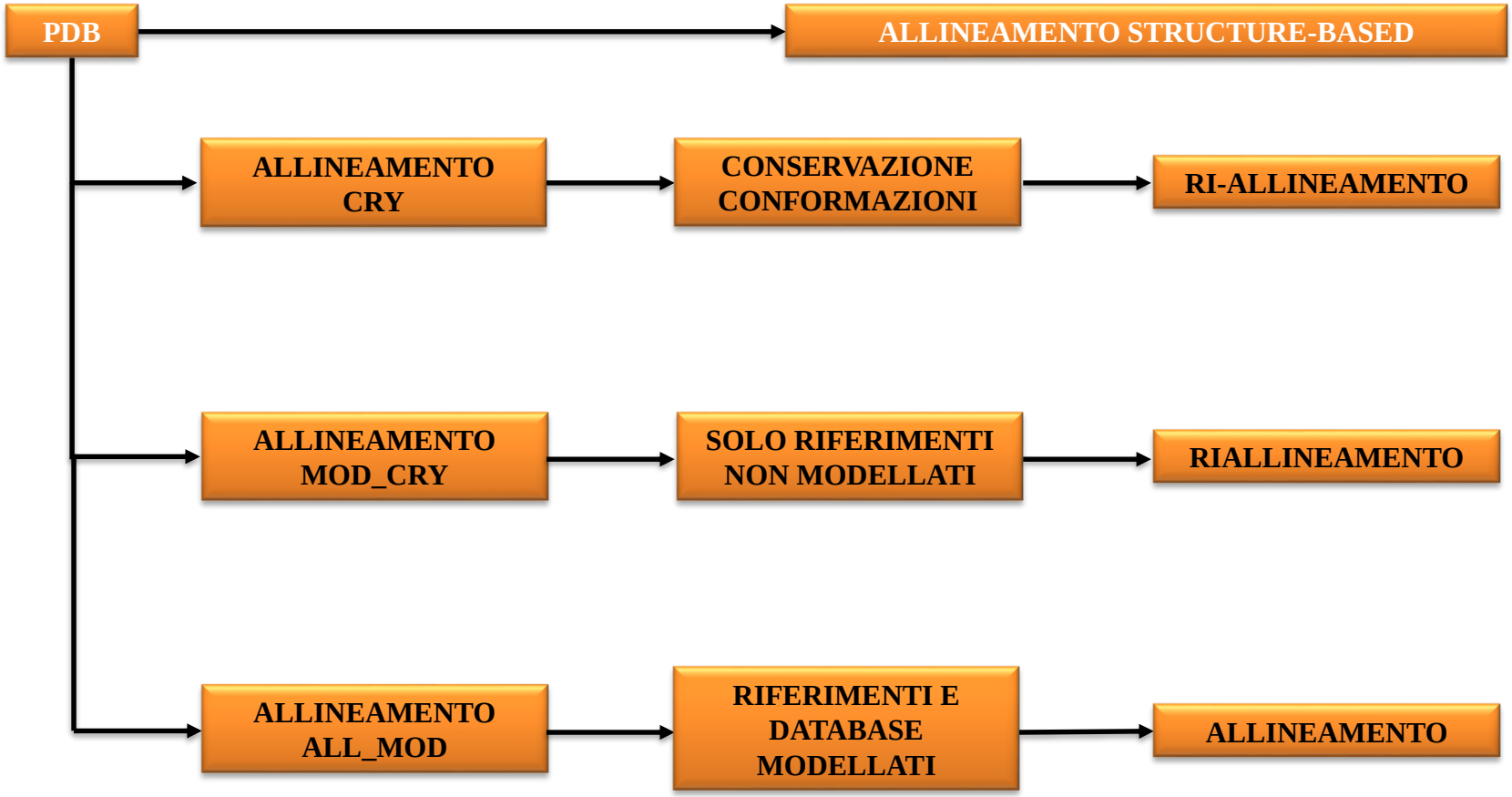
Allineamento

by www.RCMD.it

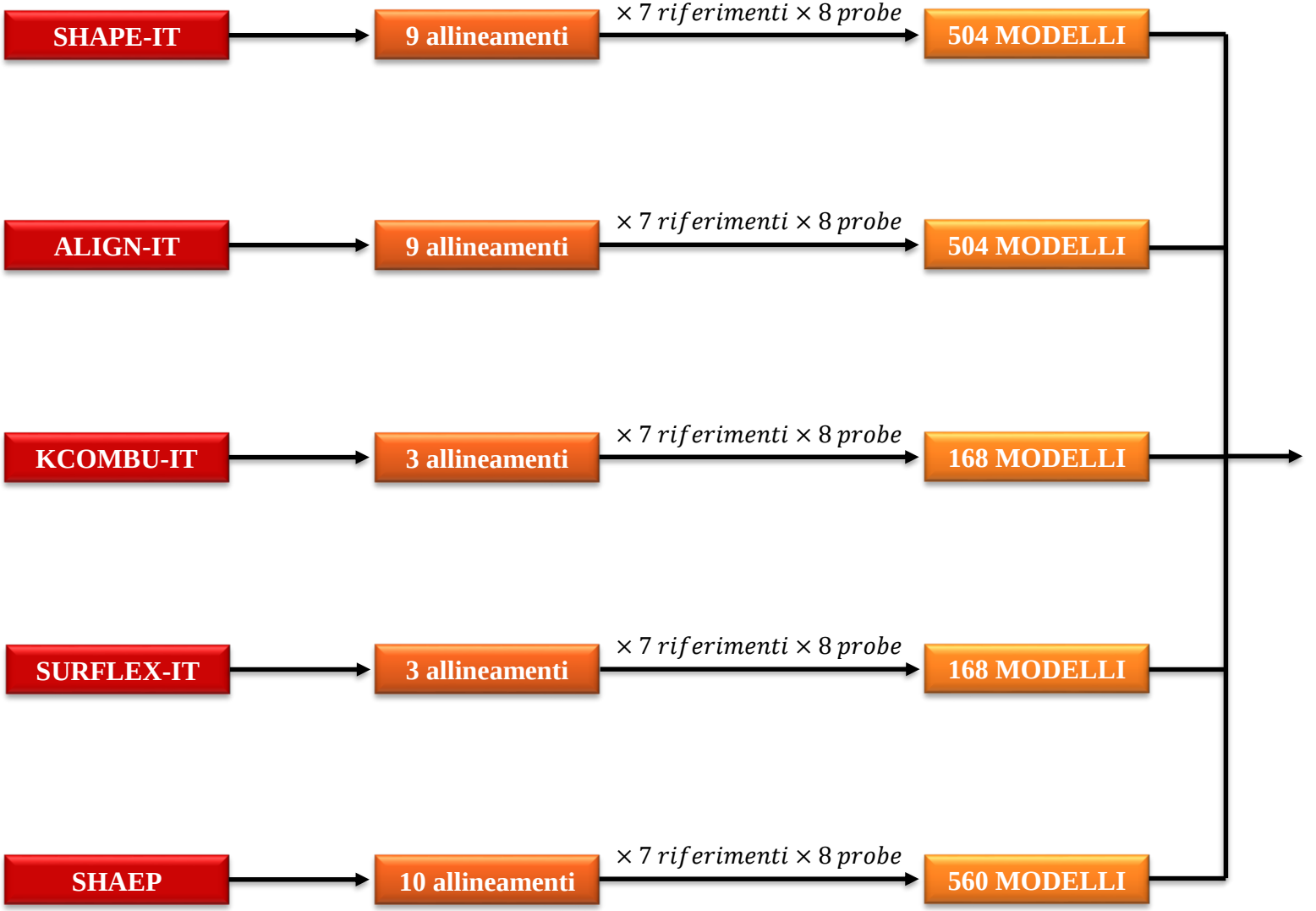


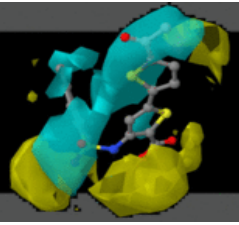
Regole di allineamento

by www.RCMD.it



Risultati Hsp90 by www.RCMD.it





Risultati VEGFR2

by www.RCMD.it

SHAPE-IT → **9 allineamenti** → $\times 7\ riferimenti \times 8\ probe$ → **504 MODELLI**

ALIGN-IT → **9 allineamenti** → $\times 7\ riferimenti \times 8\ probe$ → **504 MODELLI**

KCOMBU-IT → **3 allineamenti** → $\times 7\ riferimenti \times 8\ probe$ → **168 MODELLI**

SURFLEX-IT → **3 allineamenti** → $\times 7\ riferimenti \times 8\ probe$ → **168 MODELLI**

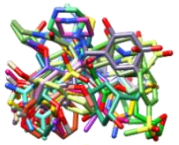
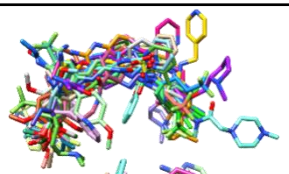
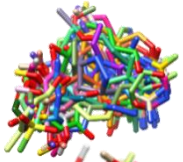
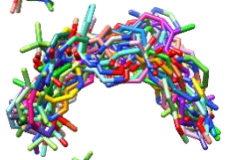
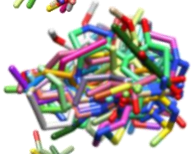
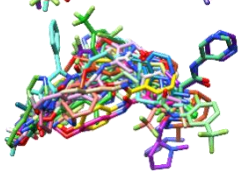
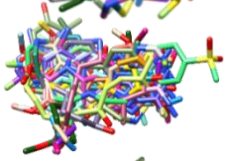
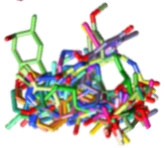
168
ALLINEAMENTI

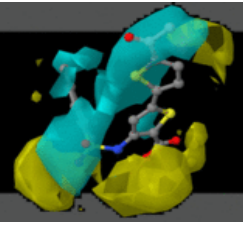
1344
MODELLI
3-D QSAR

Risultati

	N° allineamenti	N° modelli 3-D QSAR
Hsp90	238	1904
VEGFR2	168	1344

3248 MODELLI 3-D QSAR

	Hsp90	VEGFR2
ALIGN-IT All_mod		
SHAPE-IT All_mod		
SURFLEX All_mod		
KCOMBU All_mod		
SHAEP All_mod		



Confronto modelli 3-D QSAR by www.RCMD.it

HSP90

PROBE	PC	r^2	q^2 LOO	q^2 YS
A	1	0.594	0.451	-0.334
C	1	0.594	0.451	-0.317
d	3	0.657	0.342	-0.381
e	1	0.696	0.534	-0.268
HD	1	0.577	0.414	-0.308
N	**3**	**0.920**	**0.453**	**-0.556**
NA	3	0.920	0.459	-0.593
OA	3	0.922	0.460	-0.628

Modello 3-D QSAR da allineamento LB con Align-it Allmod FLEX TverskyRef sulla molecola di riferimento 1UY7, probe N

PC	r^2	SDEC	q^2	SDEP	CutOff	Zeroing	MinStd
1	0.804	0.396	0.541	0.606	5.000	0.001	0.005
2	**0.914**	**0.262**	**0.656**	**0.525**	**5.000**	**0.001**	**0.005**
3	0.946	0.208	0.655	0.526	5.000	0.001	0.005
4	0.969	0.158	0.637	0.539	5.000	0.001	0.005
5	0.986	0.106	0.595	0.570	5.000	0.001	0.0050

Numero ottimale di componenti principali: 2

VEGFR2

PROBE	PC	r^2	q^2 LOO	q^2 YS
A	5	0.986	0.667	-0.612
C	5	0.986	0.667	-0.671
d	2	0.636	0.461	-0.367
e	4	0.994	0.760	-0.593
HD	5	0.985	0.670	-1.015
N	5	0.987	0.632	-0.613
NA	5	0.988	0.631	-0.659
OA	5	0.988	0.645	-0.676

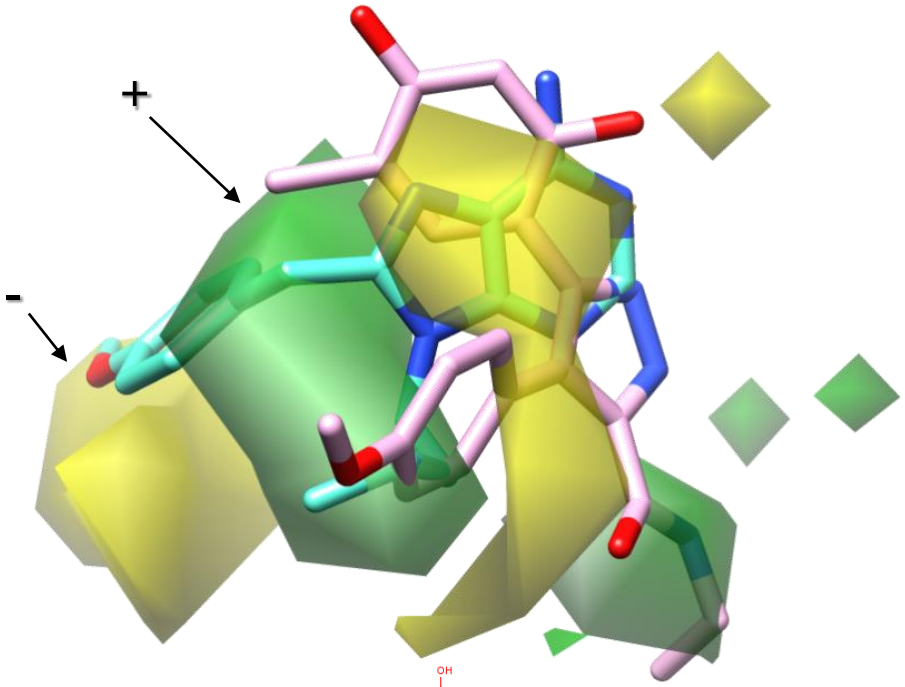
Modello 3-D QSAR da allineamento LB con Surflex Allmod sulla molecola di riferimento 3BE2, probe HD

PC	r^2	SDEC	q^2	SDEP	CutOff	Zeroing	MinStd
1	0.719	0.718	0.295	1.025	5.000	0.001	0.050
2	0.897	0.435	0.574	0.816	5.000	0.001	0.050
3	0.985	0.165	0.651	0.659	5.000	0.001	0.050
4	**0.995**	**0.093**	**0.692**	**0.587**	**5.000**	**0.001**	**0.050**
5	0.998	0.060	0.697	0.573	5.000	0.001	0.050

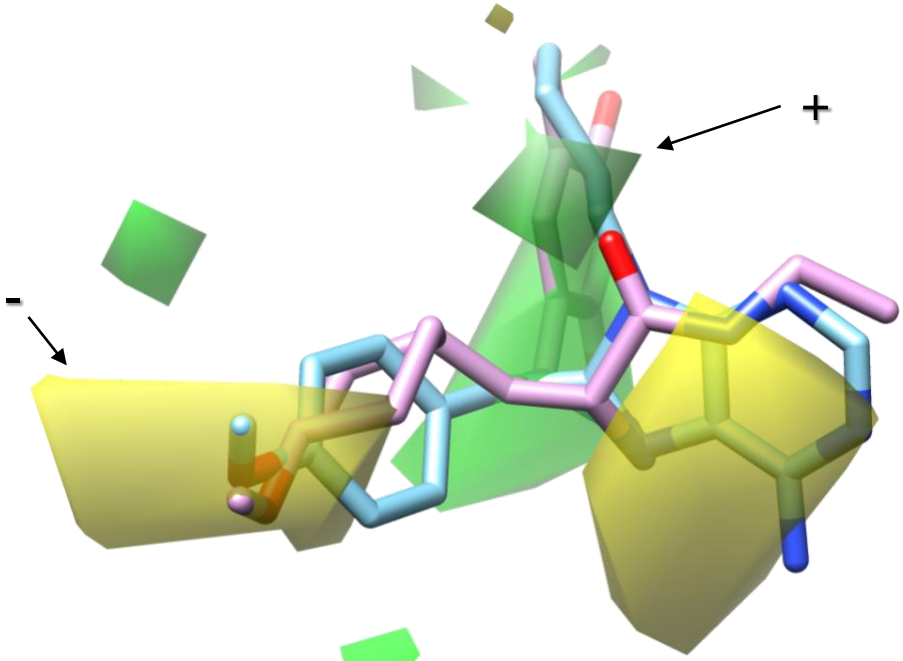
Numero ottimale di componenti principali: 4

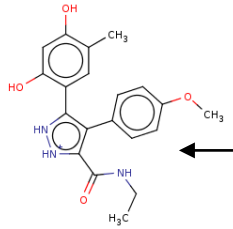
CONFRONTO MAPPE CoMFA2
PROBE N

MODELLO DI RIFERIMENTO STRUCTURE-BASED

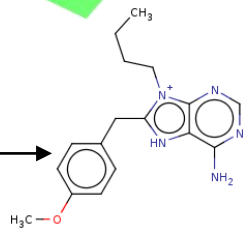


MODELLO LIGAND-BASED CON ALLINEAMENTO ALIGN-IT™ ALLMOD FLEXIBLE SULLA 1UY7



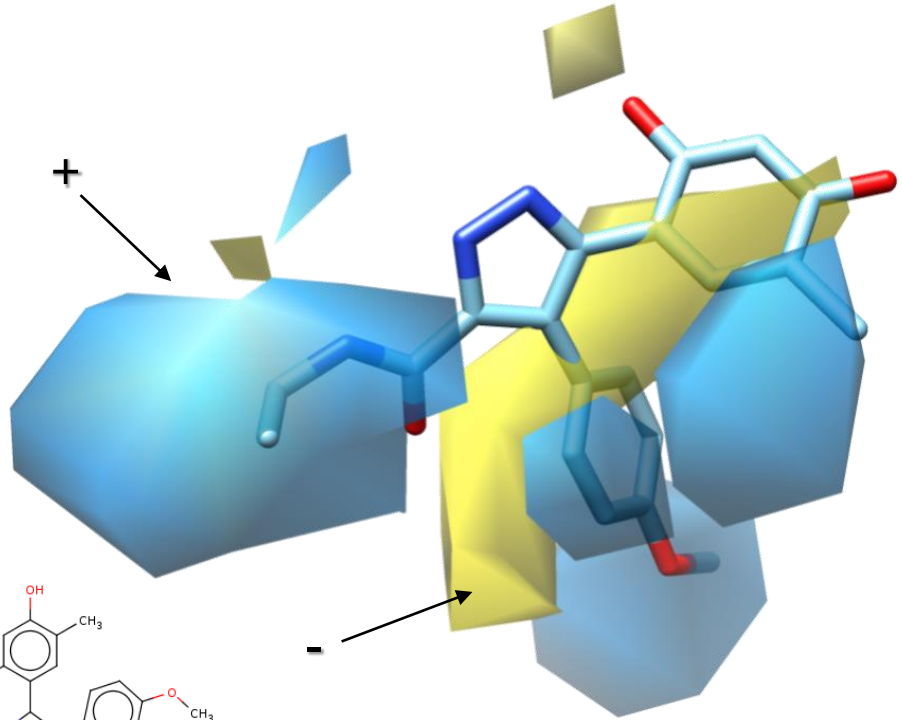


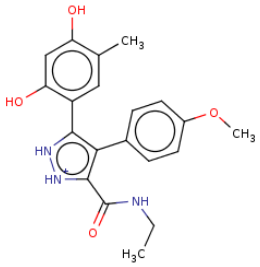
MOLECOLA PIÙ ATTIVA	MOLECOLA MENO ATTIVA
2BSM	1UY7
ROSA	CELESTE



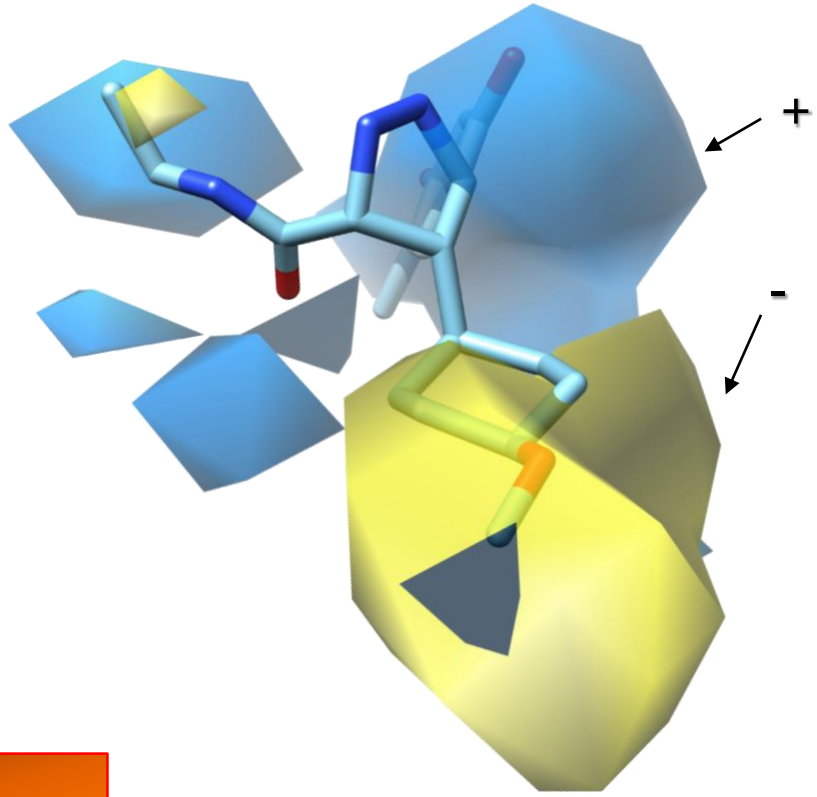
CONFRONTO MAPPE DI ACTIVITY_CONTRIBUTION PROBE N

MODELLO DI RIFERIMENTO





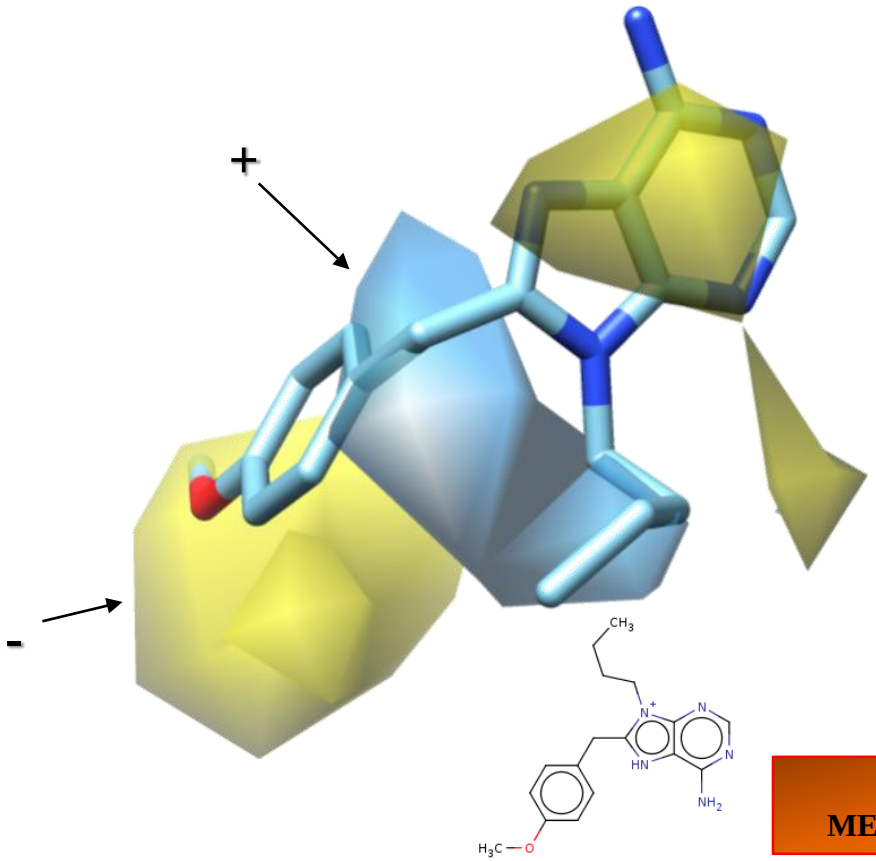
MODELLO LIGAND-BASED CON ALLINEAMENTO ALIGN-IT™ ALLMOD FLEXIBLE SULLA 1UY7



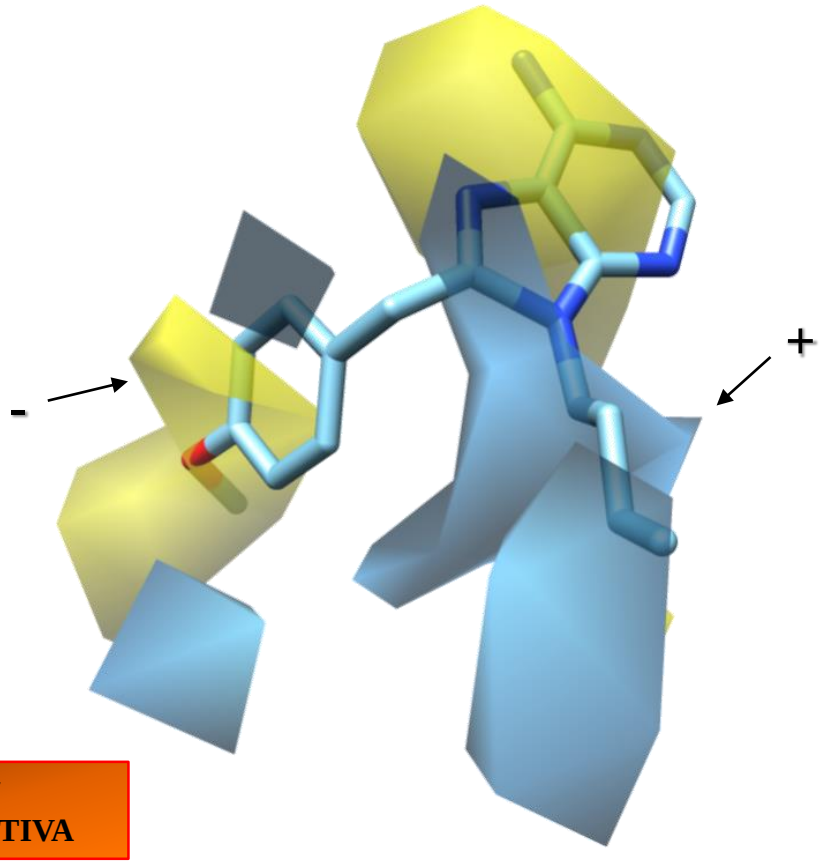
2BSM
PIÙ ATTIVA

CONFRONTO MAPPE DI ACTIVITY_CONTRIBUTION
PROBE N

MODELLO DI RIFERIMENTO



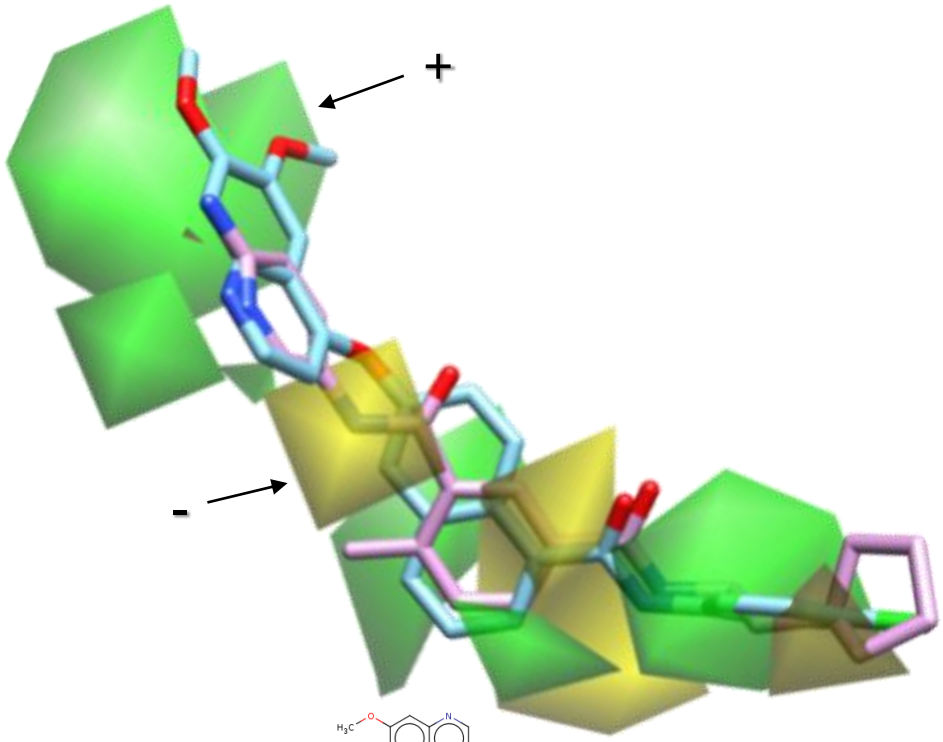
MODELLO LIGAND-BASED CON ALLINEAMENTO ALIGN-IT™ ALLMOD FLEXIBLE SULLA 1UY7



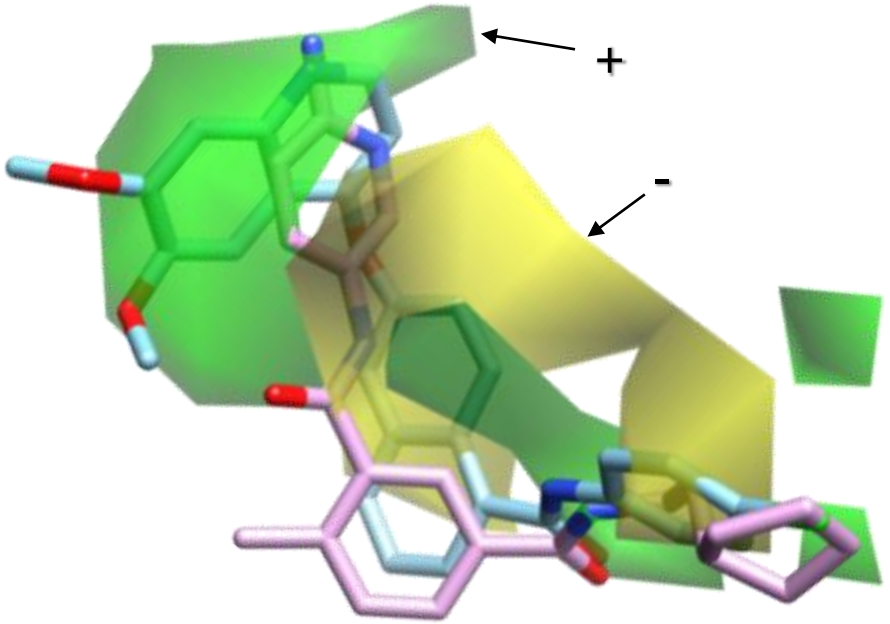
1UY7
MENO ATTIVA

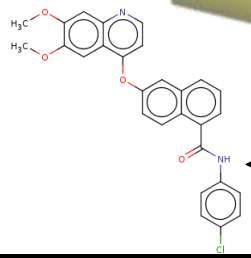
CONFRONTO MAPPE CoMFA2
PROBE HD

MODELLO DI RIFERIMENTO STRUCTURE-BASED

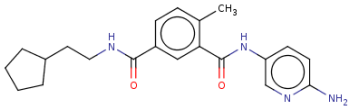


MODELLO LIGAND-BASED CON ALLINEAMENTO SURFLEX-SIM ALLMOD SULLA 3BE2



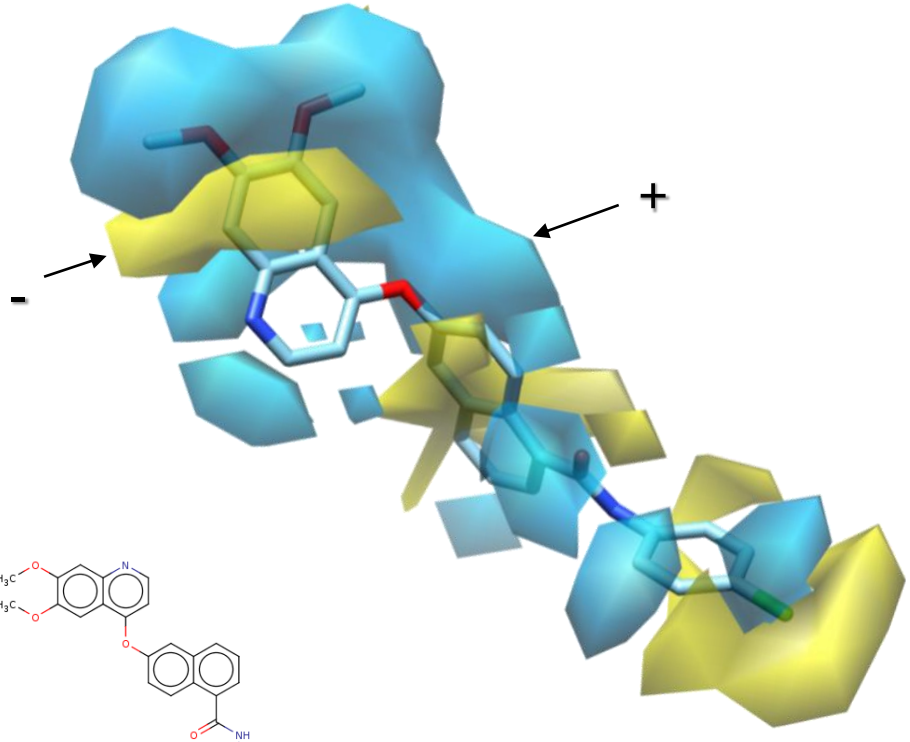


MOLECOLA PIÙ ATTIVA	MOLECOLA MENO ATTIVA
3B8Q_A **CELESTE**	3CPB_A **ROSA**

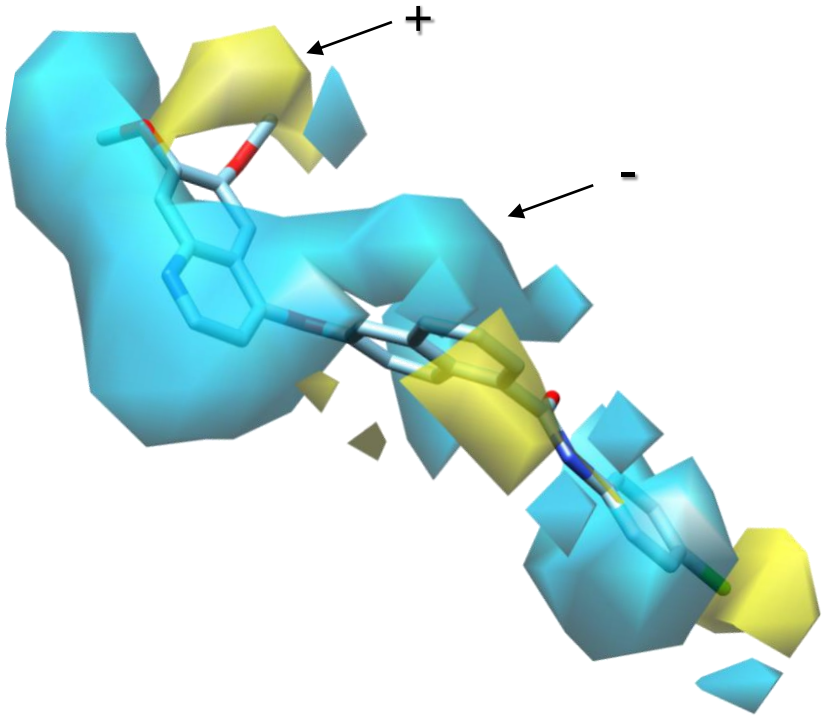


CONFRONTO MAPPE DI ACTIVITY_CONTRIBUTION
PROBE HD

MODELLO DI RIFERIMENTO



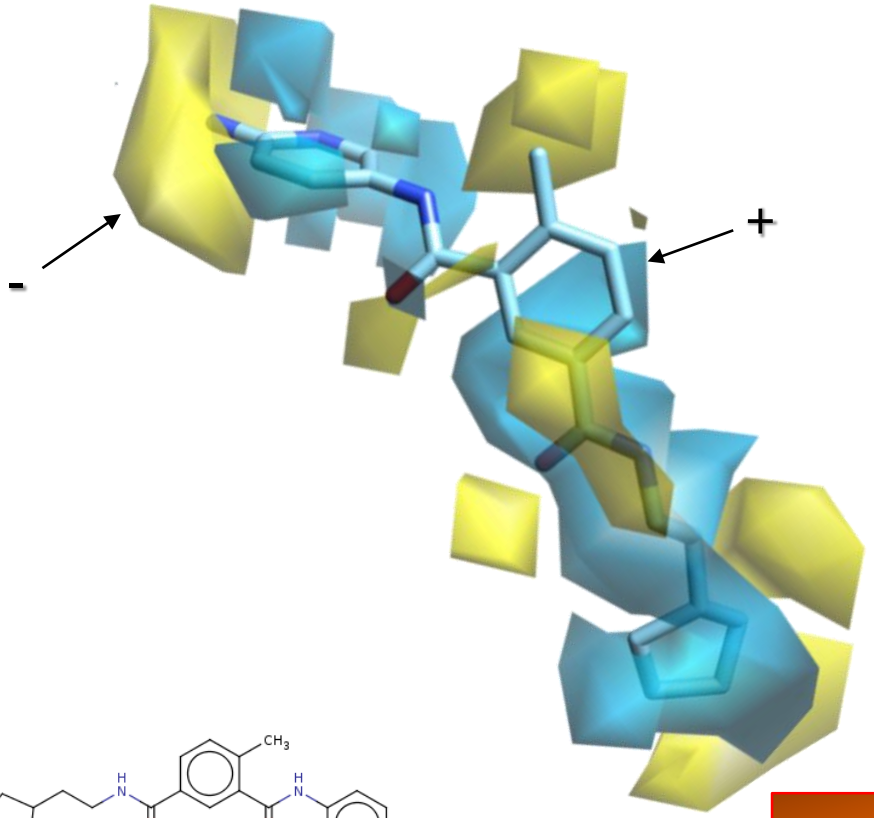
MODELLO LIGAND-BASED CON ALLINEAMENTO ALIGN-IT™ ALLMOD FLEXIBLE SULLA 1UY7



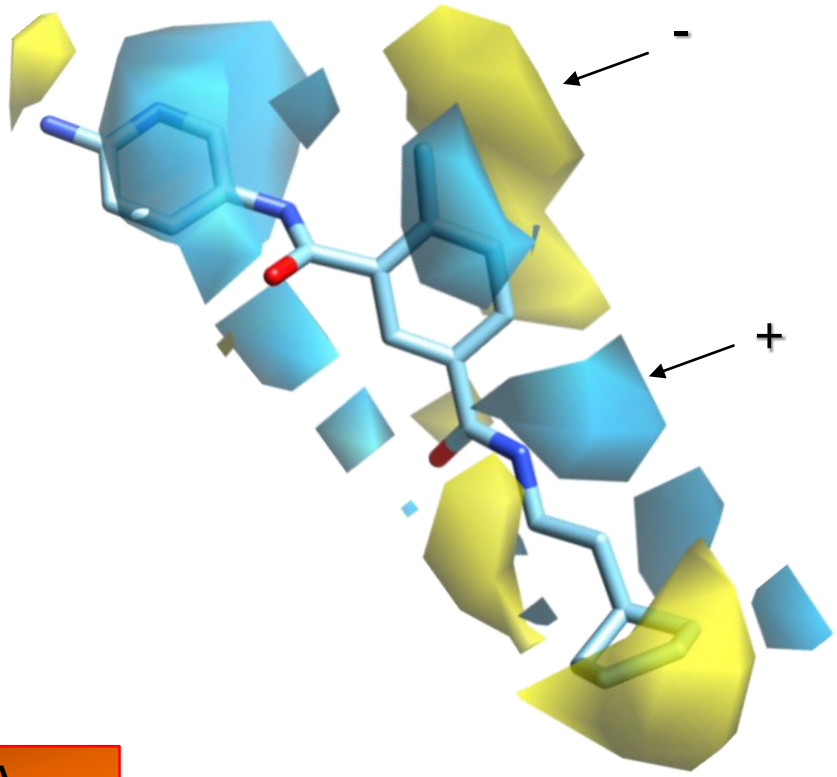
3B8Q_A
PIÙ ATTIVA

CONFRONTO MAPPE DI ACTIVITY_CONTRIBUTION
PROBE HD

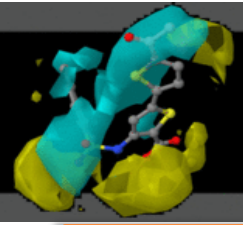
MODELLO DI RIFERIMENTO



MODELLO LIGAND-BASED CON ALLINEAMENTO ALIGN-IT™ ALLMOD FLEXIBLE SULLA 1UY7



3CPB_A
MENO ATTIVA



Conclusioni

by www.RCMD.it

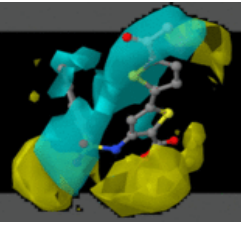
Modelli 3-D QSAR ligand-based riescono a fornire le stesse informazioni dei modelli 3-D QSAR structure-based di riferimento

Validazione del protocollo per la generazione automatica dei modelli 3-D QSAR tramite il programma di multi-allineamento QSAlign.py e il metodo Py_3-D_QSAR (versione Python e migliorata di 3-D QSAutogrid/R)

Dimostrazione dell'affidabilità dei modelli 3-D QSAR nei casi in cui uno studio ligand-based è l'unico approccio

OBIETTIVI FUTURI

- aumentare il numero di molecole di riferimento e quindi di allineare tutte le strutture tra di loro reciprocamente
- allestimento di un database di modelli 3-D QSAR per predire l'attività biologica di molecole non saggiate, all'interno di un portale web



Grazie per l'attenzione

" We try to combine hard core molecular biology,

hard core computer sciences and hard core statistics because

that's what it will take to answer the questions of the 21st century."

Guillaume Filion, Centre of Genomic Regulation ,Spain