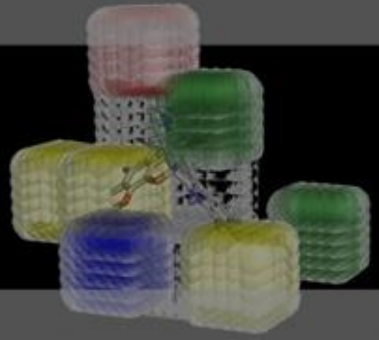
www.rcmd.it
Rome Center For Molecular Design



Sapienza Università di Roma
facoltà di farmacia

Sviluppo, validazione ed applicazione di una procedura alternativa per la costruzione e l'analisi grafica di relazioni quantitative struttura-attività di tipo tridimensionale (3-D QSAR).

Relatore: Prof. Rino Ragno

Laureando:Flavio Ballante

CoMFA

The Lattice Model: A General Paradigm for Shape-Related Structure/Activity Correlation

Cramer, R.D., and Milne, M., *Abstracts ACS Meeting*, Honolulu, 1979, COMP 44.

SIAM J. Sci. and Stat. Comput. / Volume 5 / Issue 3

The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses

S. Wold, A. Ruhe, H. Wold, and W. J. Dunn, III

SIAM J. Sci. and Stat. Comput. Volume 5, Issue 3, pp. 735-743

•(1988)

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

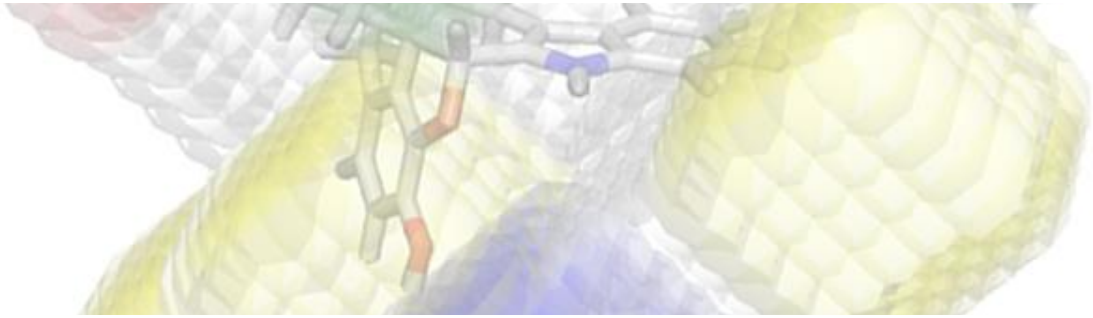
Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins

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*

GRID/GOLPE

Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S., **Generating Optimal Linear Pls Estimations (Golpe) - an Advanced Chemometric Tool for Handling 3d-Qsar Problems**.
Quant Struct-Act Rel 1993, 12, (1), 9-20.



•(1994)

J. Med. Chem. **1994,** *37,* 2589−2601

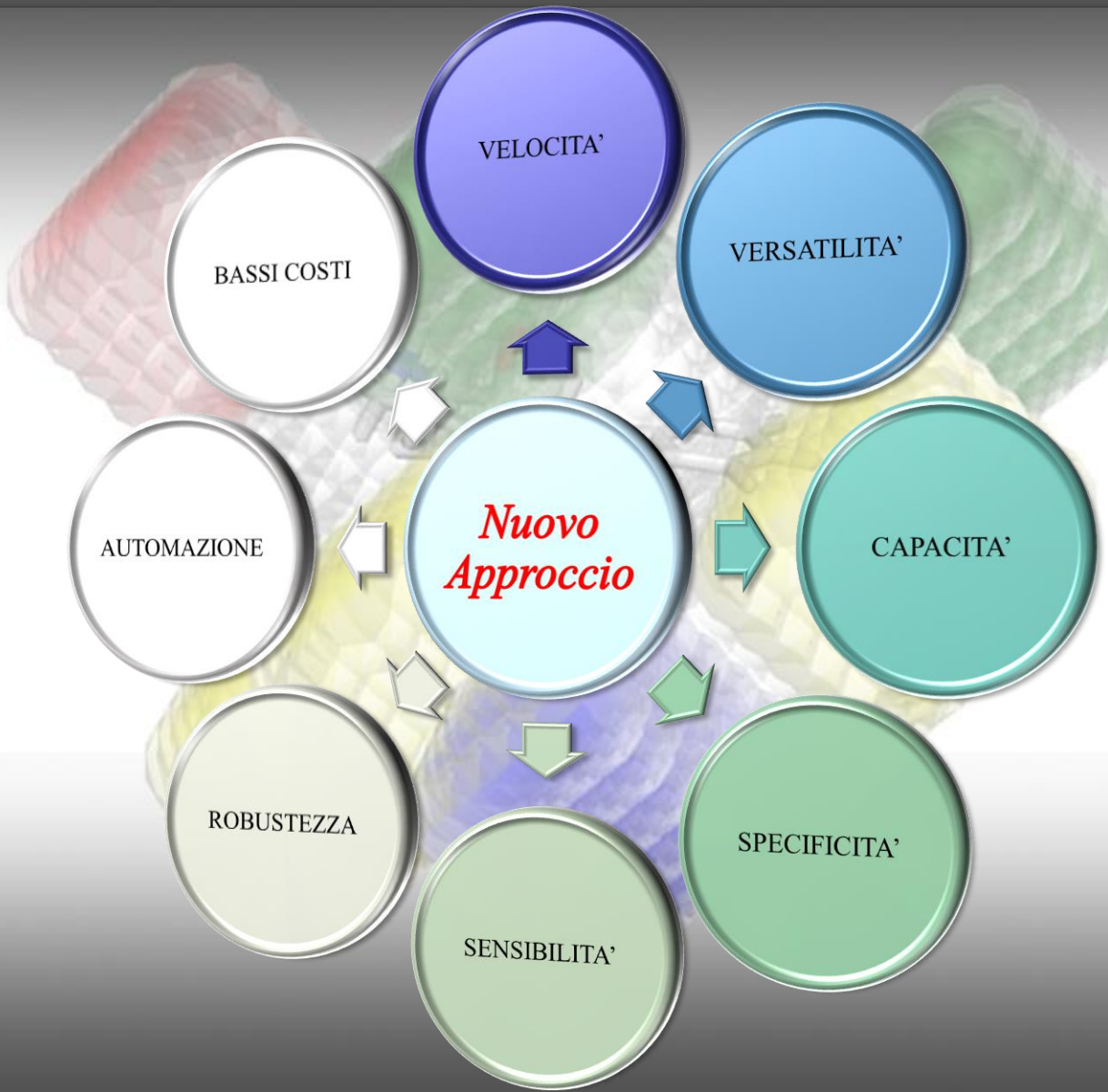
Comparative Molecular Field Analysis Using GRID Force-Field and GOLPE Variable Selection Methods in a Study of Inhibitors of Glycogen Phosphorylase *b*

Gabriele Cruciani*,† and Kimberly A. Watson‡

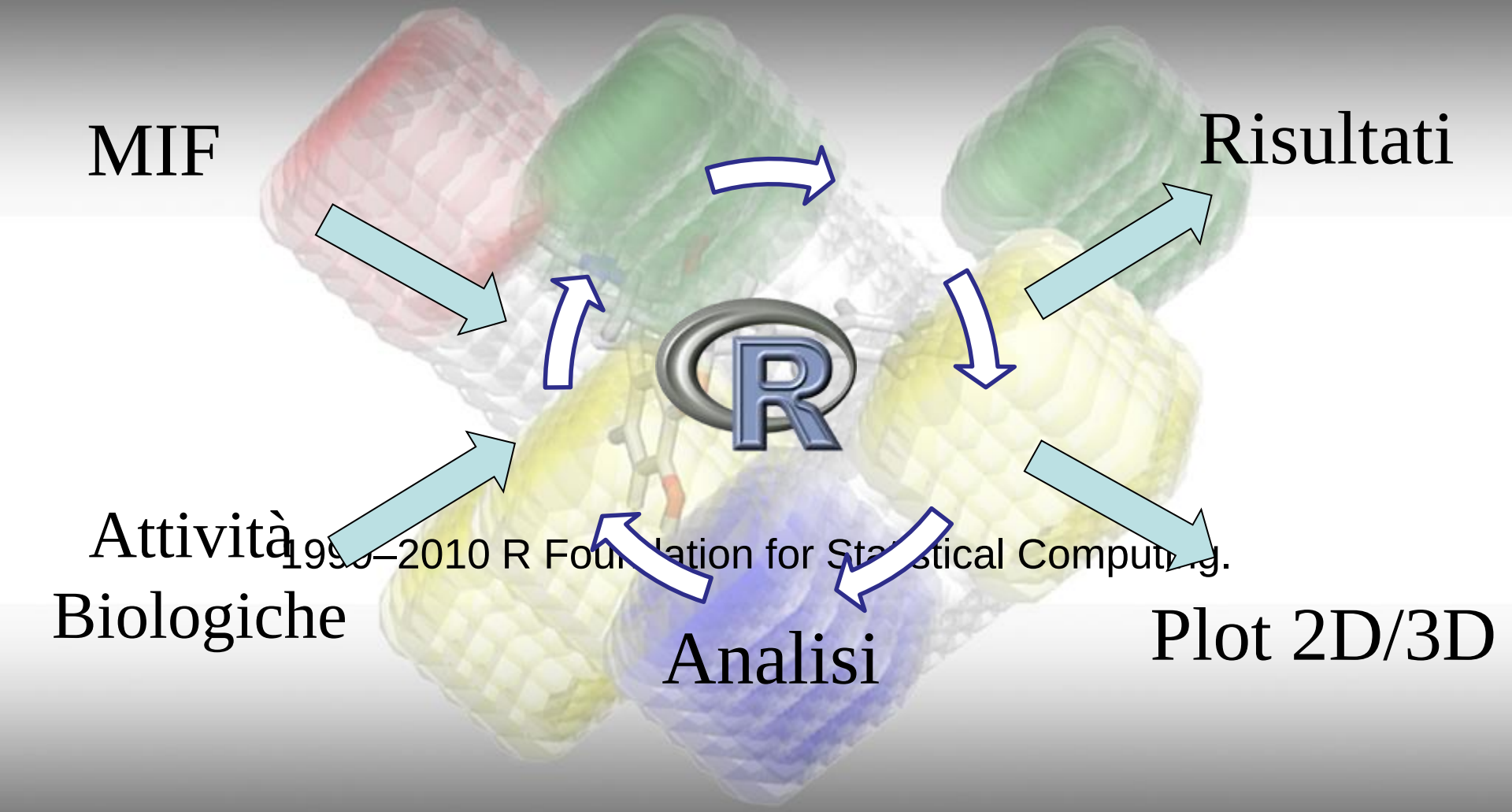
Department of Chemistry, University of Perugia, Via Elce di Sotto, 8, 06100 Perugia, Italy, and Laboratory of Molecular Biophysics, University of Oxford, South Parks Road, OX1 3QU, Oxford, England

Sviluppo 3-D QSAR

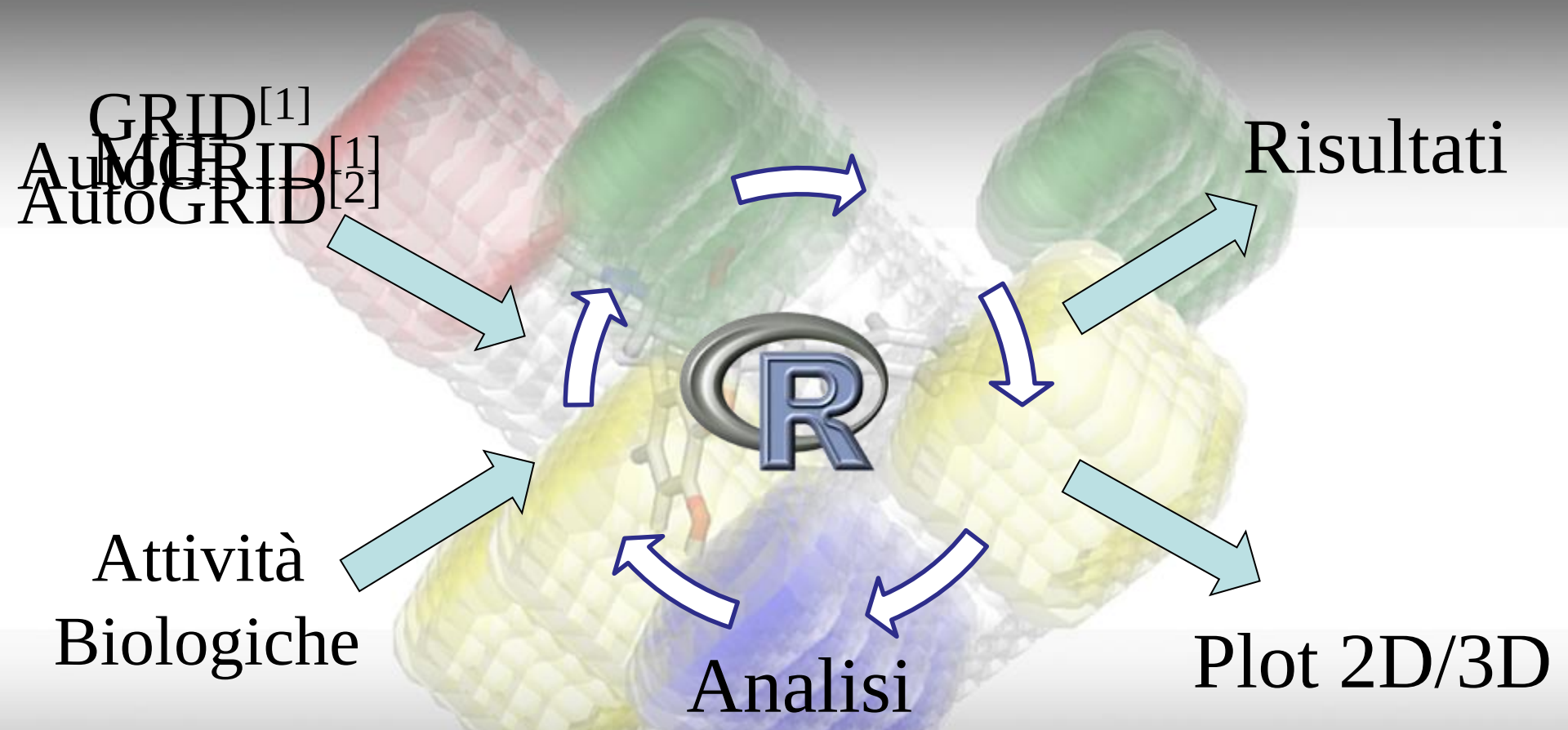
Requisiti di Analisi



Architettura R Procedurale



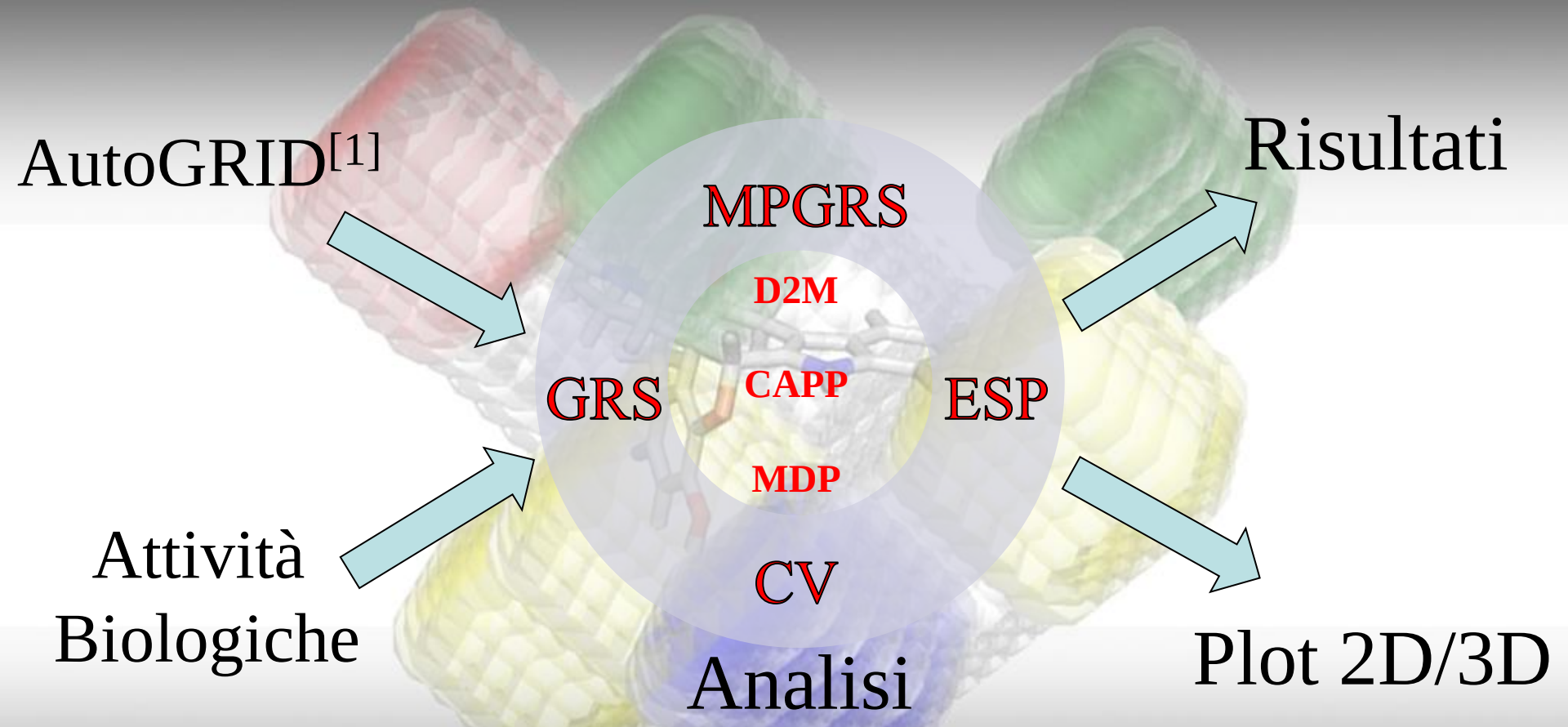
Architettura Procedurale



1. G. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, **Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.** *J. Comput. Chem.* ***1998, 19, (14), 1639-1662.***

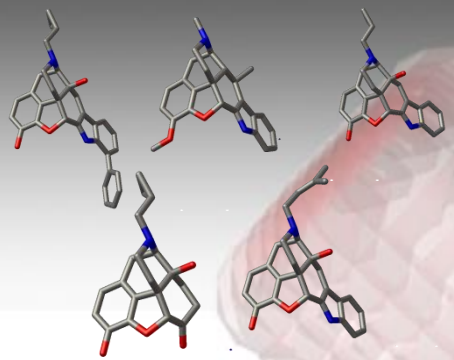
2. G. M. Morris, D. S. G., R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, **Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.** *J. Comput. Chem.* ***1998, 19, (14), 1639-1662.***

Funzioni di Calcolo

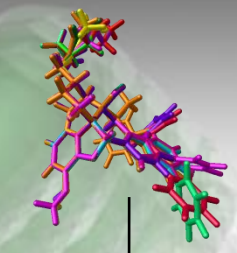


1. G. M. Morris, D. S. G., R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, **Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function.** *J. Comput. Chem.* ***1998, 19, (14), 1639-1662.***

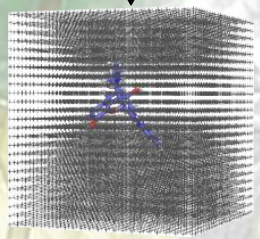
Procedura Globale



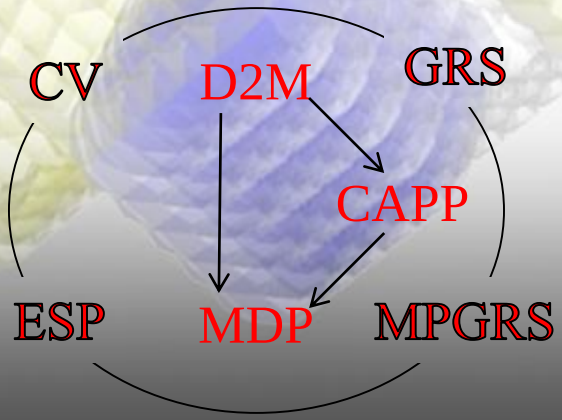
Serie Molecolare di studio
(TRAINING SET)



ALLINEAMENTO
Structure/Ligand Based

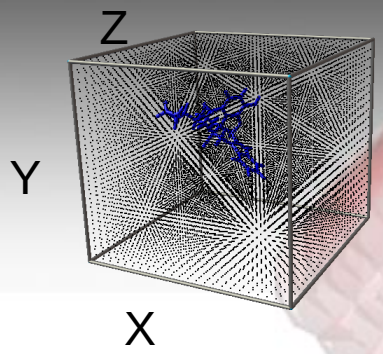


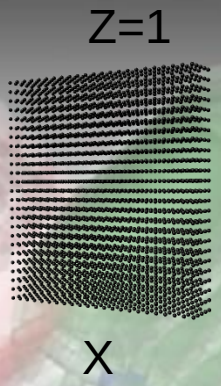
GENERAZIONE MIF
(AutoGRID)

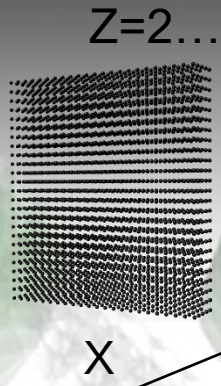


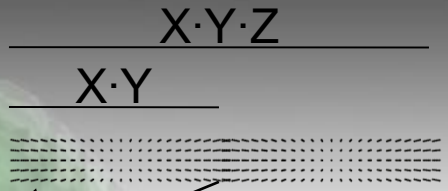
Pacchetto D2M

"Data to Model"





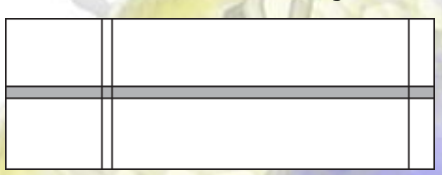




Attività Biologiche

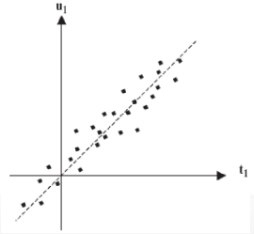
Coordinate x,y,z

Composto x



Analisi PLS

(r^2, SDEC)

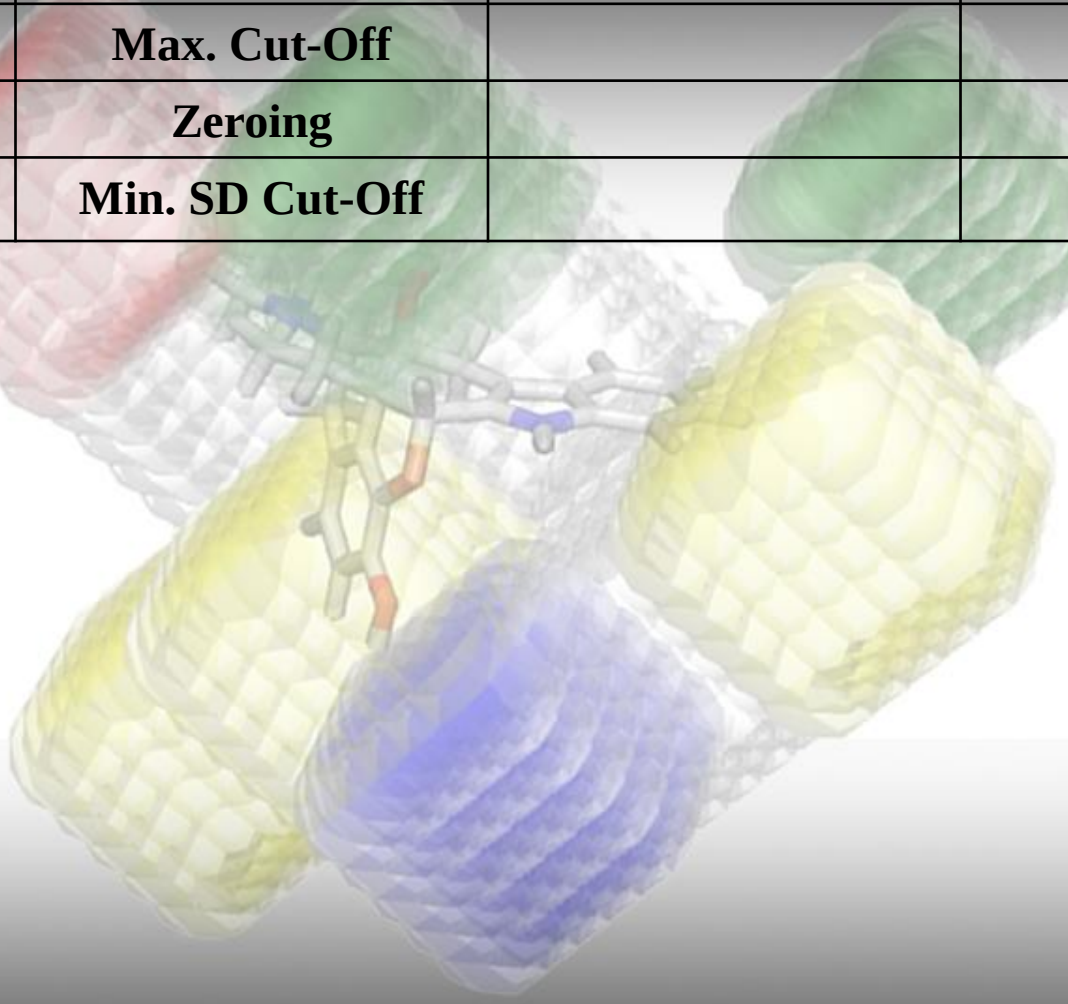


Modello 3-DQSAR "RAW"

Pacchetto CAPP

“Combinatorial Analysis of Pretreatment Parameters”

Valore Minimo	Parametro	Valore Massimo	Passo
	Max. Cut-Off		
	Zeroing		
	Min. SD Cut-Off		



Pacchetto CAPP

"Combinatorial Analysis of Pretreatment Parameters"

Valore Minimo	Parametro	Valore Massimo	Passo
0	**Max. Cut-Off**	*1.00*	*0.50*
0.10	**Zeroing**	*0.20*	*0.05*
0.05	**Min. SD Cut-Off**	*0.06*	*0.01*

N° Comb.	Cut-Off	Zeroing	SD
1	*0.00*	*0.01*	*0.05*
2	*0.05*	*0.01*	*0.05*
3	*1.00*	*0.01*	*0.05*
4	*0.00*	*0.15*	*0.05*
5	*0.05*	*0.15*	*0.05*
6	*1.00*	*0.15*	*0.05*
7	*0.00*	*0.02*	*0.05*
8	*0.05*	*0.02*	*0.05*
9	*1.00*	*0.02*	*0.06*
10	*0.00*	*0.01*	*0.06*
11	*0.05*	*0.01*	*0.06*
12	*1.00*	*0.01*	*0.06*
13	*0.00*	*0.15*	*0.06*
14	*0.05*	*0.15*	*0.06*
15	*1.00*	*0.15*	*0.06*
16	*0.00*	*0.02*	*0.06*
17	*0.05*	*0.02*	*0.06*
18	*1.00*	*0.02*	*0.06*

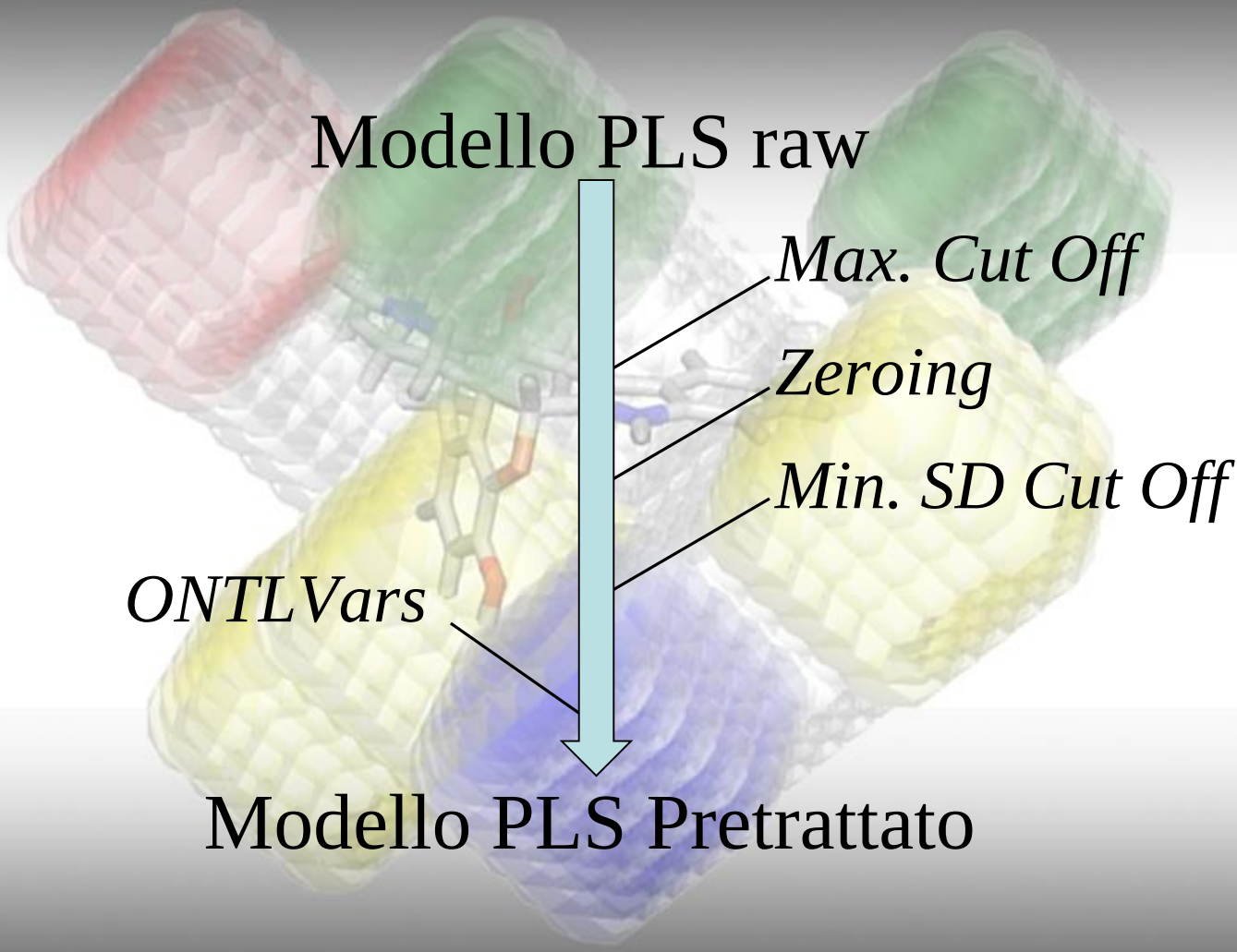
Confronto tra i modelli generati (q^2, SDEP)

↓

Combinazione migliore

Pacchetto MDP

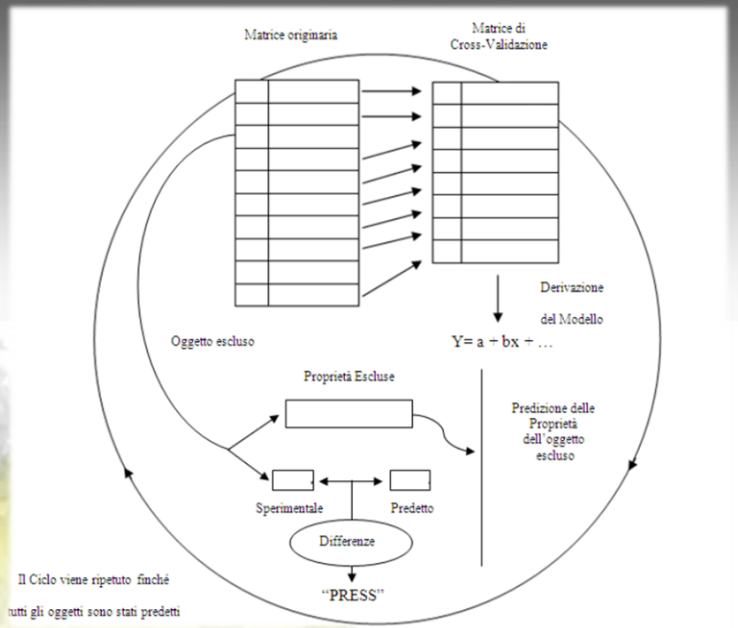
"Model Data Pretreatment"



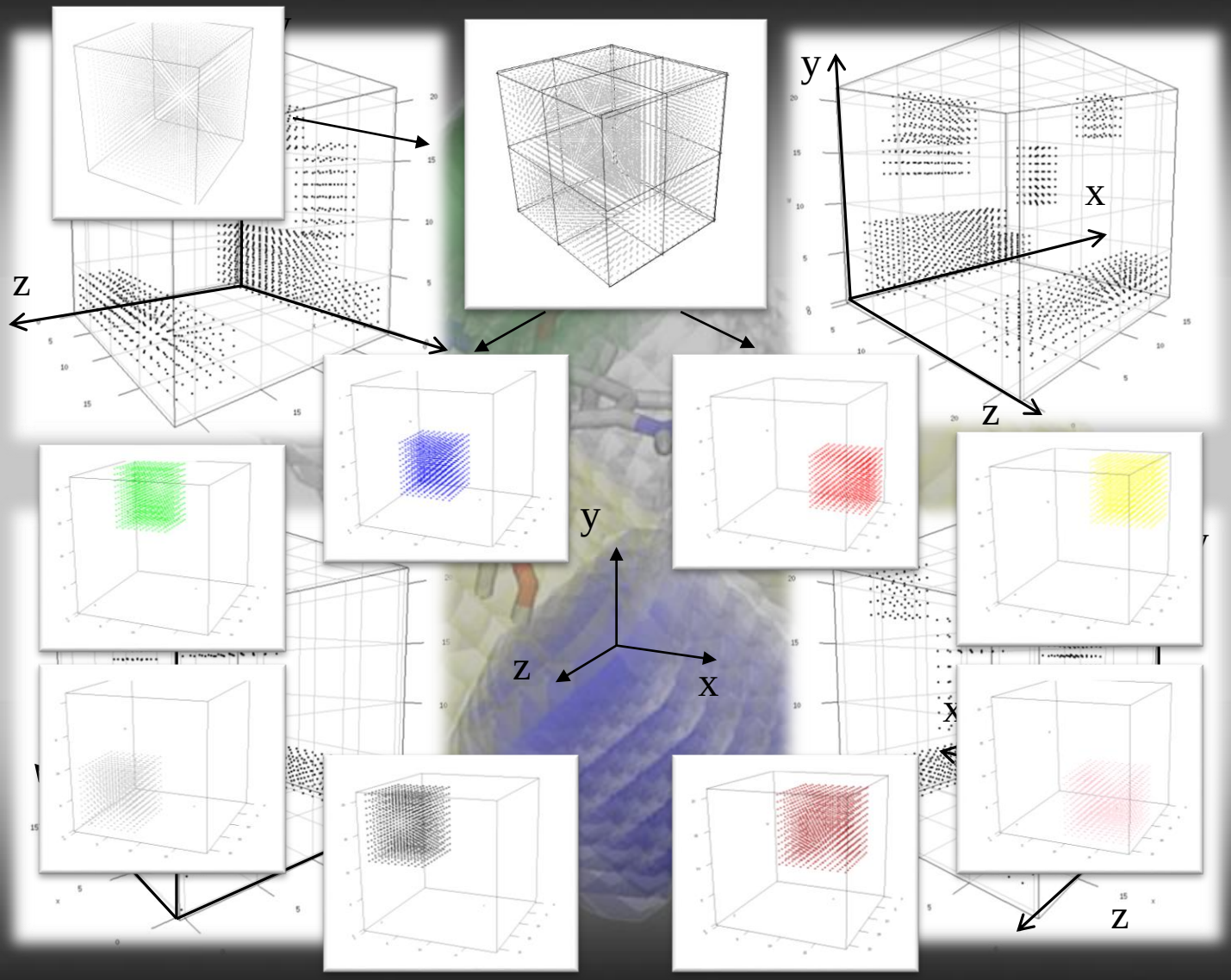
Pacchetto CV
“Cross-Validation”

Cross-Validazione:

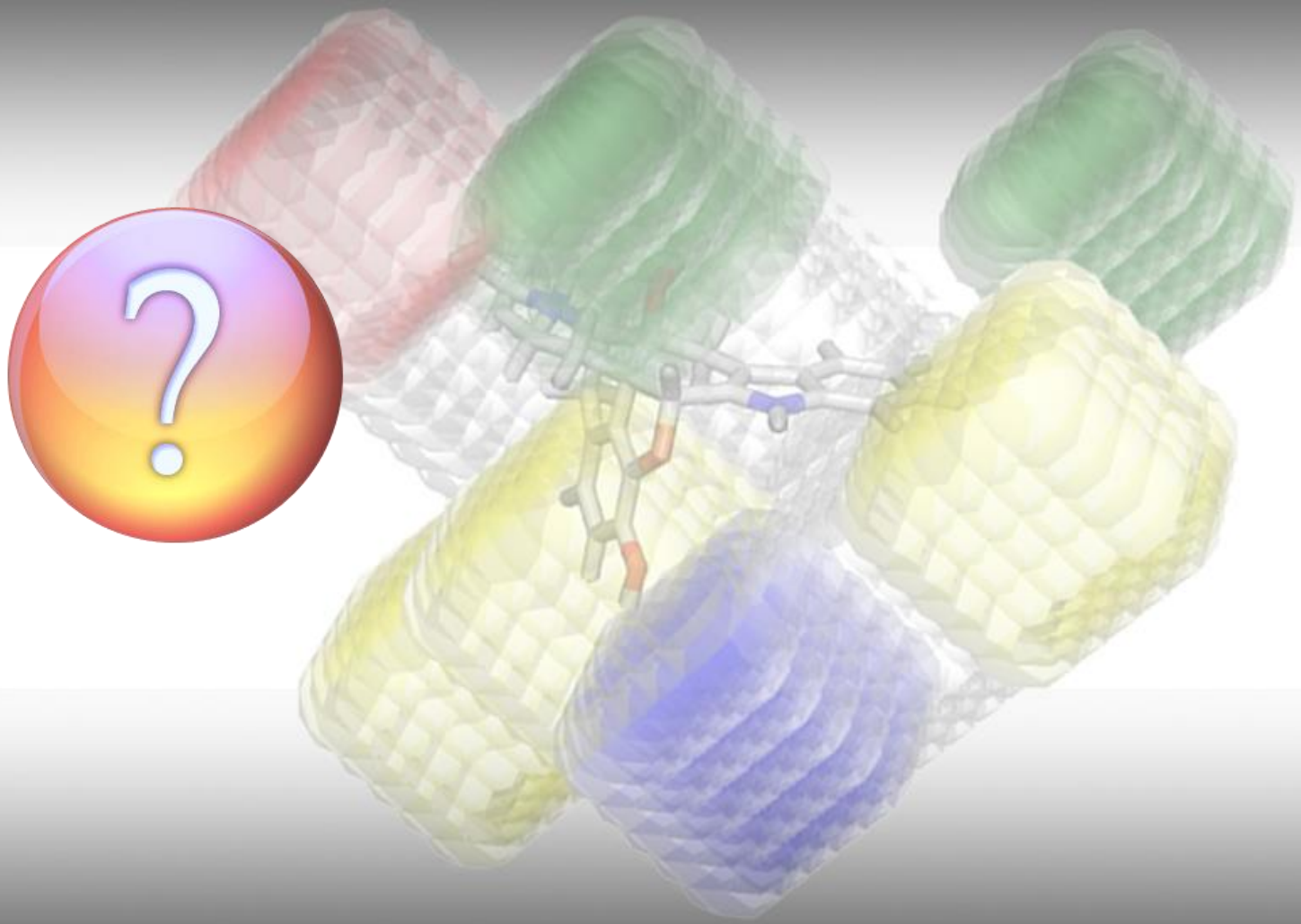
- LOOCV:Leave One Out
- LSOCV:Leave Some Out
- K-FOLDCV:Leave Groups Out
  - **Random**
  - **Consecutive**
  - **Interleaved**
- MCCV: Monte-Carlo



Pacchetto GRS
“Guided Region Selection”



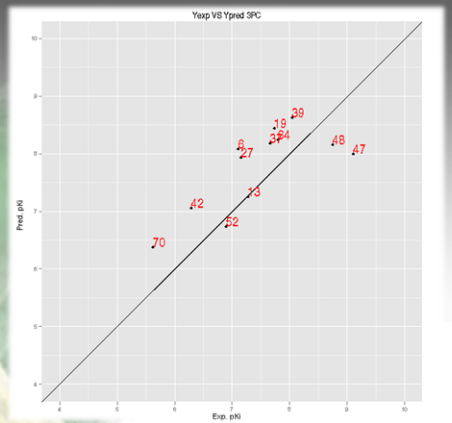
Pacchetto ESP
“External Set Prediction”



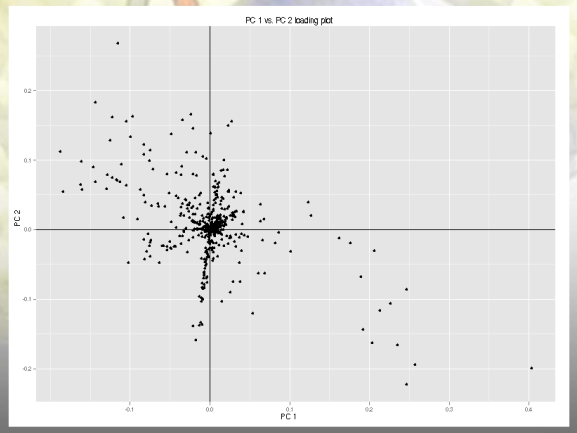
Rappresentazioni grafiche: PLOT 2-D

Score Plot

Inner Correlation Plot



Y_{exp} vs Y_{pred}

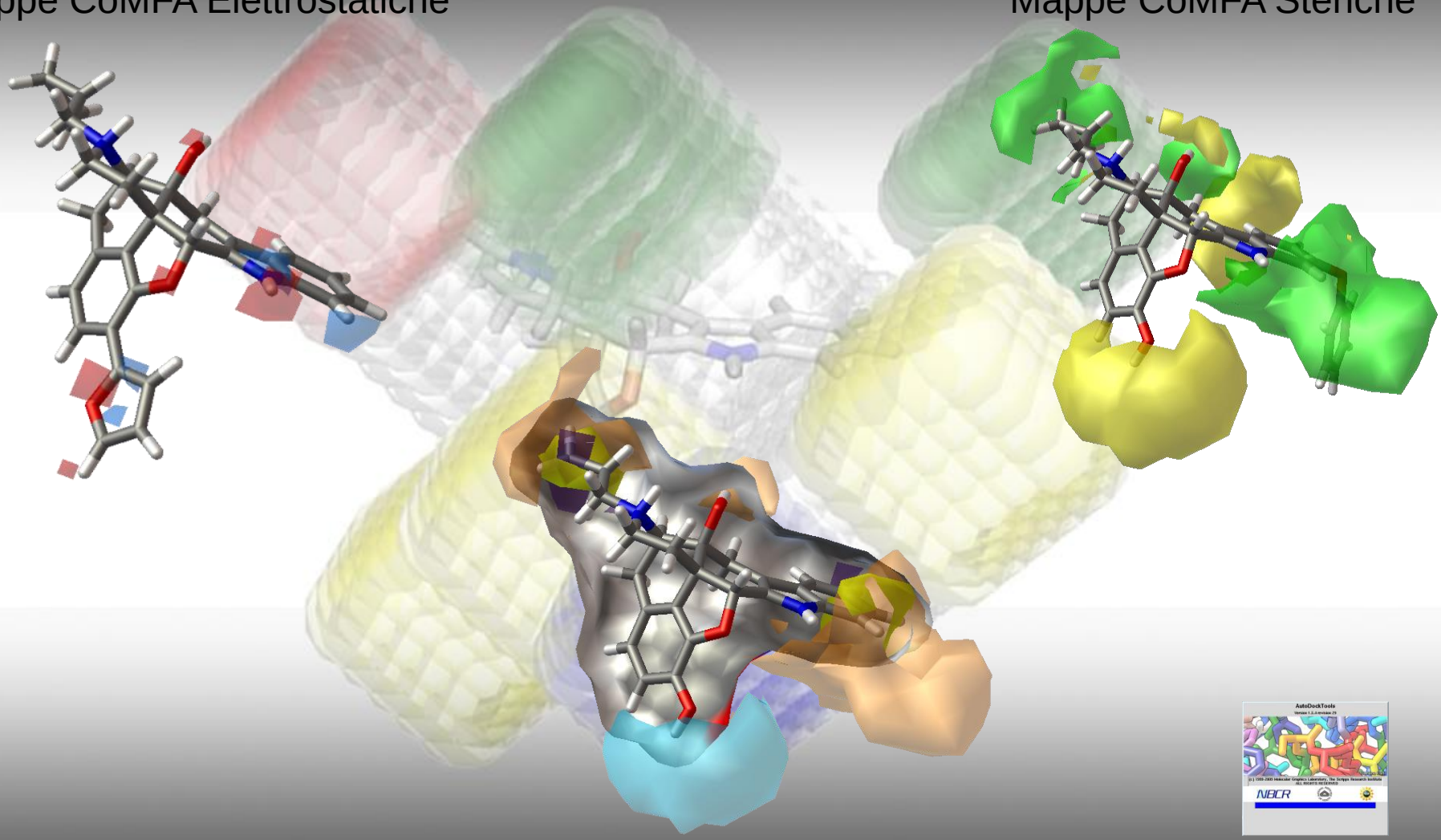


Loading Plot

Rappresentazioni grafiche: PLOT 3-D

Mappe CoMFA Elettrostatiche

Mappe CoMFA Steriche



PLS Coefficient + Act. Contr. Plot

Applicazioni

- Antagonisti recettori oppioidi

1620 *J. Med. Chem.* **2005,** *48,* 1620−1629

3D-QSAR Comparative Molecular Field Analysis on Opioid Receptor Antagonists: Pooling Data from Different Studies

Youyi Peng, Susan M. Keenan, Qiang Zhang, Vladyslav Kholodovych, and William J. Welsh*

Department of Pharmacology and the Informatics Institute of UMDNJ, University of Medicine & Dentistry of New Jersey−Robert Wood Johnson Medical School (UMDNJ−RWJMS), Piscataway, New Jersey 08854

- Inibitori della RNA-polimerasi NS5B dell'HCV

J. Chem. Inf. Model. **XXXX,** *xxx,* 000 A

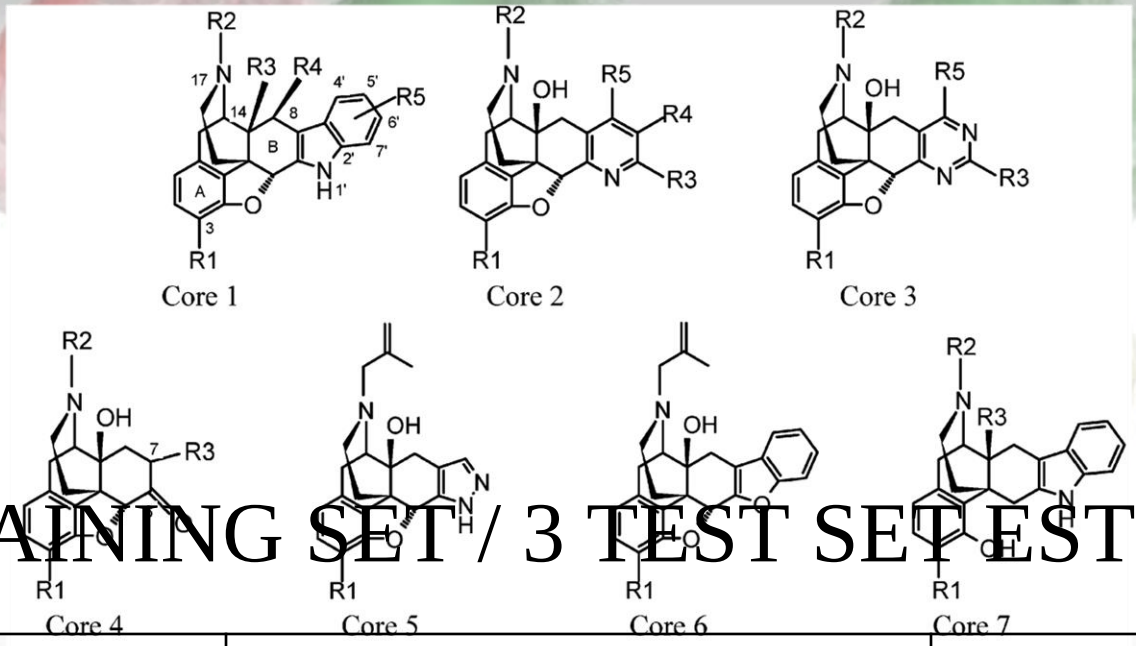
Combining 3-D Quantitative Structure−Activity Relationship with Ligand Based and Structure Based Alignment Procedures for *in Silico* Screening of New Hepatitis C Virus NS5B Polymerase Inhibitors

Ira Musmuca,[†] Antonia Caroli,[†] Antonello Mai,[†] Neerja Kaushik-Basu,[‡] Payal Arora,[‡] and Rino Ragno*[,†]

Istituto Pasteur-Fondazione Cenci Bolognetti, Dipartimento di Chimica e Tecnologie del Farmaco, Sapienza Università di Roma, P. le A. Moro 5, 00185, Rome, Italy and Department of Biochemistry and Molecular Biology, UMDNJ-New Jersey Medical School, 185 South Orange Avenue, Newark, NJ

Antagonisti dei Recettori Oppioidi

74 composti

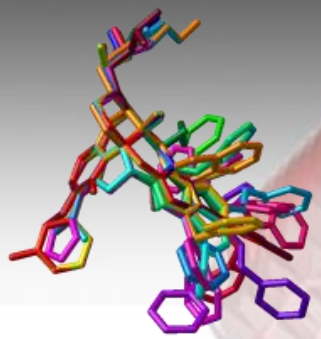


3 TRAINING SET / 3 TEST SET ESTERNI

RECETTORE	N° compd TRAINING SET	N° compd TEST SET
δ	61	12
μ	48	10
κ	49	10

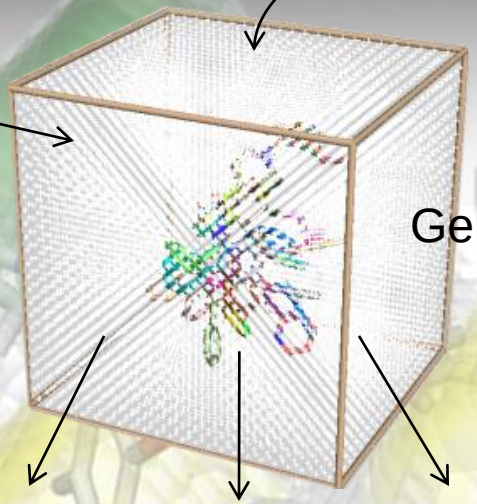
7 Strutture di base (Cores)

Antagonisti dei Recettori Oppioidi



Allineamento
LIGAND-BASED

8 Probe: A, C, HD, OA, NA, N, e, d



Generazione MIF

Costruzione Modelli

δ **μ** **κ**

A, C, HD, OA, NA, N, e, d

8 PROBE × 3 SETS = 24 MODELLI PLS

Antagonisti dei Recettori Oppioidi

Non pretrattati “RAW”:

- LOOCV
- L2OCV
- KF5CV (100 iter)
- MCCV(4:1, 200 iter)

Analisi CAPP:

Valore Minimo	Parametro	Valore Massimo	Passo
0	Max. Cut-Off	50	5.0
0	Zeroing	0.1	0.01
0	Min. SD Cut-Off	5	0.05

CV

24 MODELLI OTTIMIZZATI

Risultati statistici dei modelli 3-D QSAR

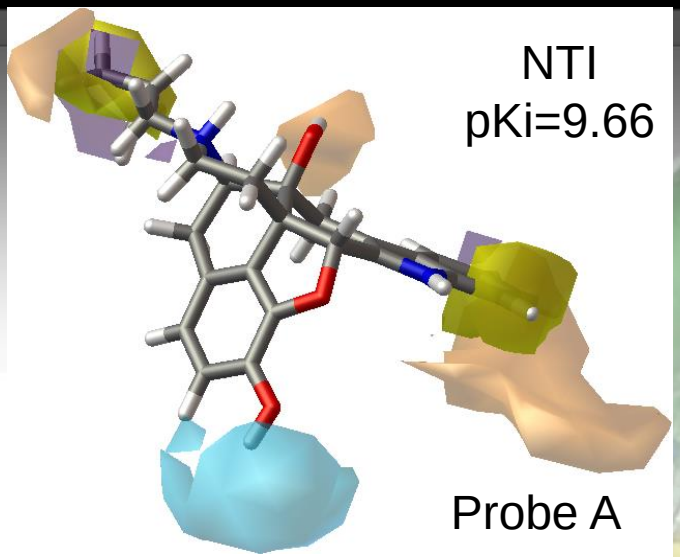
24 MODELLI 3-D QSAR OTTIMIZZATI

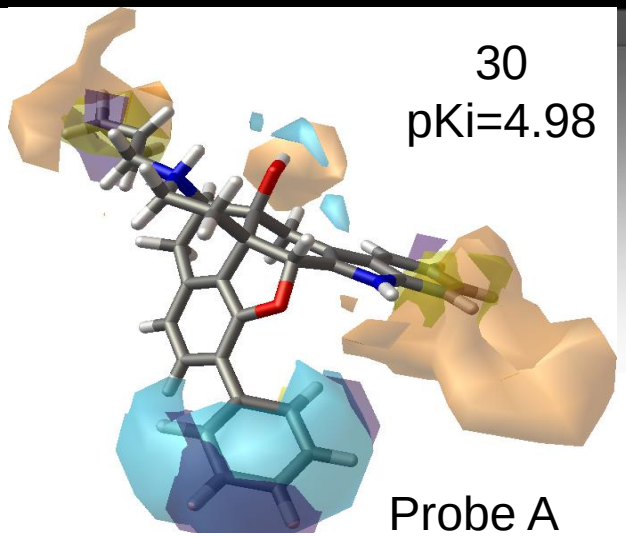
MODELLO δ PRETRATTATO						
PROBE	PC	r^2	q^2 LOO	q^2 L2O	q^2 KF5	q^2 MCCV
A	3	0.81	0.68	0.67	0.65	0.66
C	3	0.82	0.68	0.67	0.65	0.68
HD	3	0.82	0.67	0.66	0.64	0.68
NA	3	0.82	0.69	0.67	0.66	0.66
N	3	0.81	0.68	0.67	0.65	0.63
OA	3	0.82	0.69	0.68	0.65	0.69
e	3	0.60	0.40	0.41	0.35	0.32
d	3	0.65	0.51			

MODELLO μ PRETRATTATO						
PROBE	PC	r^2	q^2 LOO	q^2 L2O	q^2 KF5	q^2 MCCV
A	4	0.90	0.73	0.73	0.67	0.66
C	4	0.90	0.73	0.73	0.68	0.68
HD	4	0.89	0.72	0.71	0.67	0.68
NA	3	0.84	0.56	0.55	0.52	0.49
N	4	0.89	0.73	0.73	0.69	0.69
OA	4	0.89	0.73	0.73	0.69	0.77
e	1	0.25	0.12	0.13	0.11	0.09
				0.50	0.43	0.49

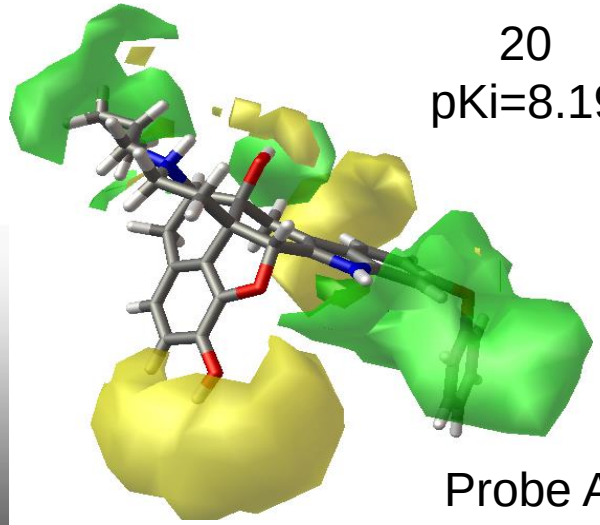
MODELLO κ PRETRATTATO						
PROBE	PC	r^2	q^2 LOO	q^2 L2O	q^2 KF5	q^2 MCCV
A	2	0.68	0.39	0.37	0.32	0.40
C	2	0.68	0.39	0.38	0.32	0.20
HD	4	0.78	0.58	0.58	0.52	0.46
NA	2	0.70	0.42	0.40	0.32	0.26
N	2	0.68	0.40	0.38	0.32	0.35
OA	2	0.69	0.41	0.39	0.33	0.39
e	1	0.15	-0.02	-0.03	-0.05	0.00
d	3	0.50	0.28	0.27	0.26	0.26

Mappe di correlazione recettore δ

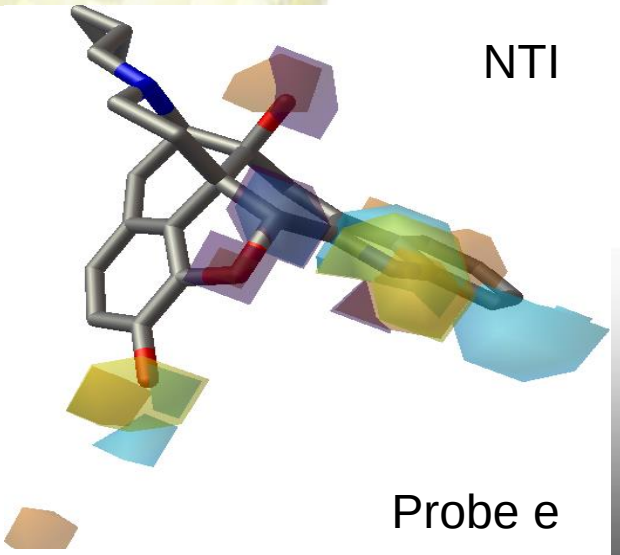




PLS-Coeff.+Act.Contr. Plot

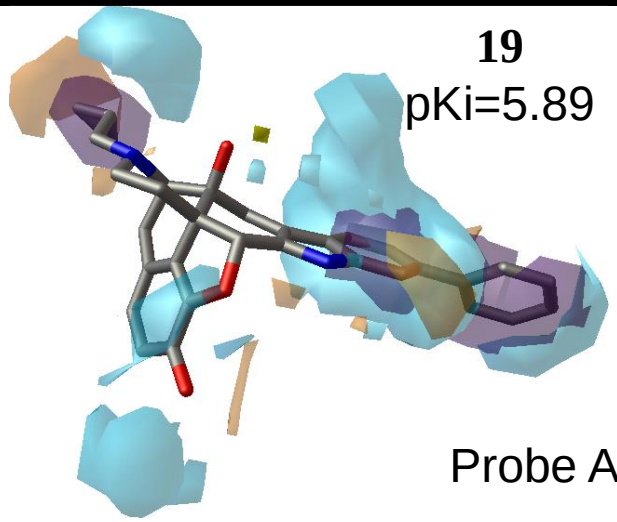


Mappa CoMFA sterica

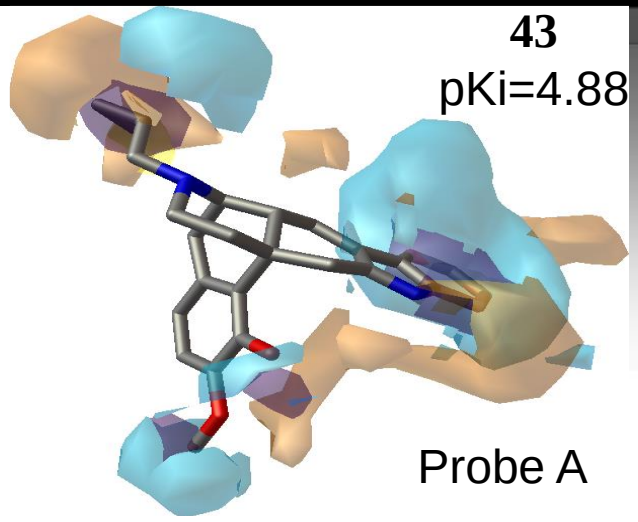


PLS-Coeff.+Act.Contr. Plot

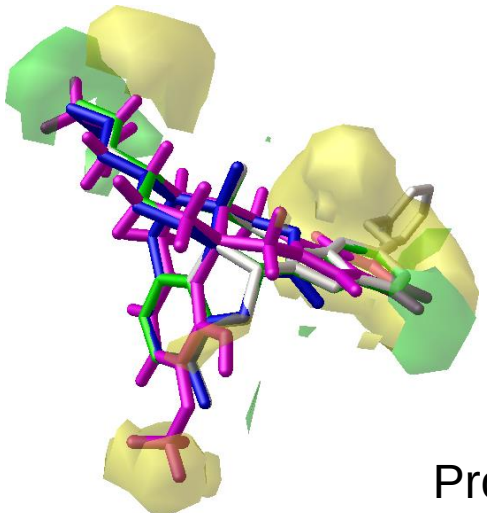
Mappe di correlazione recettore μ



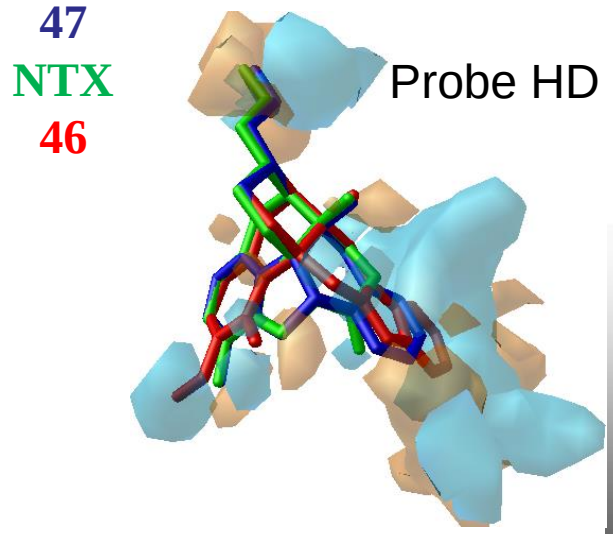
PLS-Coeff.+Act.Contr. Plot



PLS-Coeff.+Act.Contr. Plot

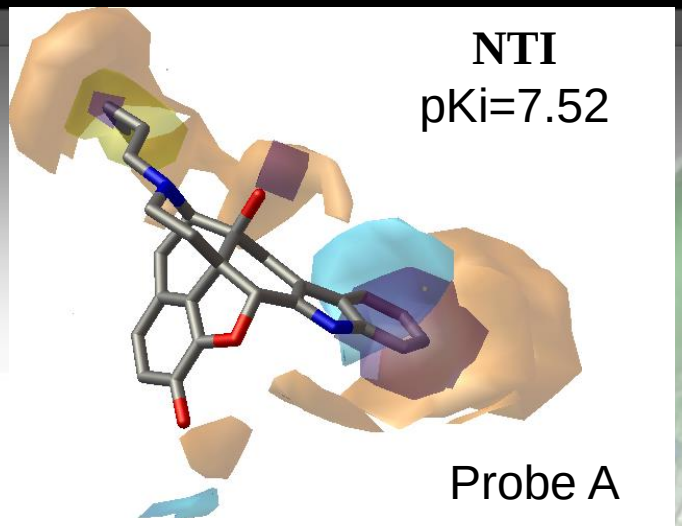


Mappa CoMFA sterica

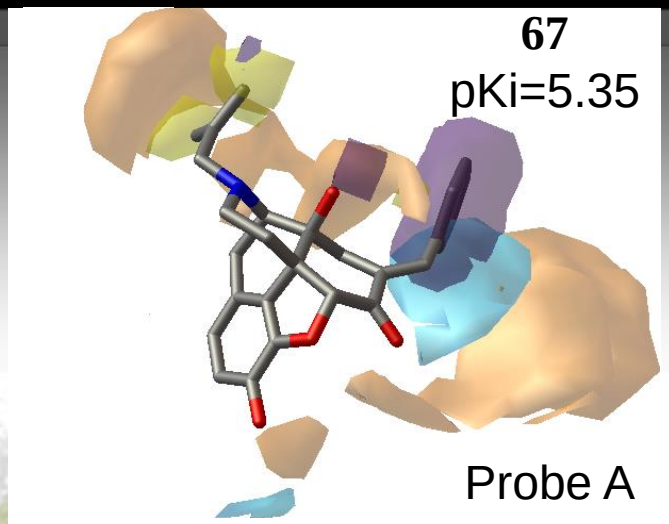


PLS-Coeff. Plot

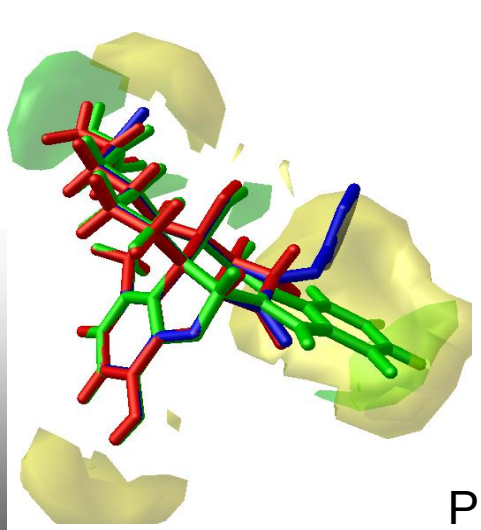
Mappe di correlazione recettore κ



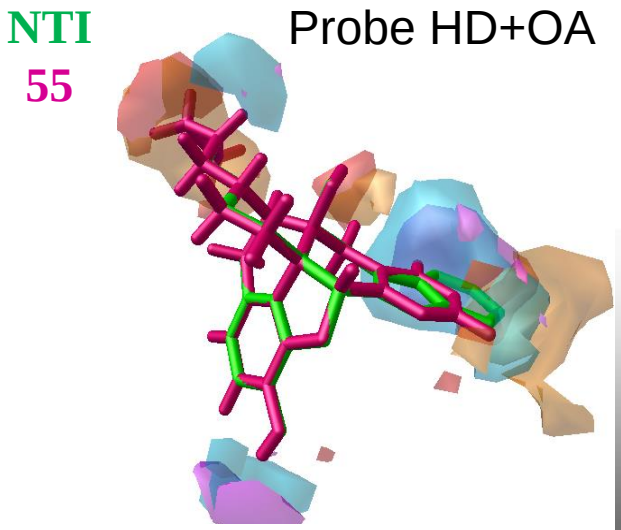
PLS-Coeff.+Act.Contr. Plot



PLS-Coeff.+Act.Contr. Plot



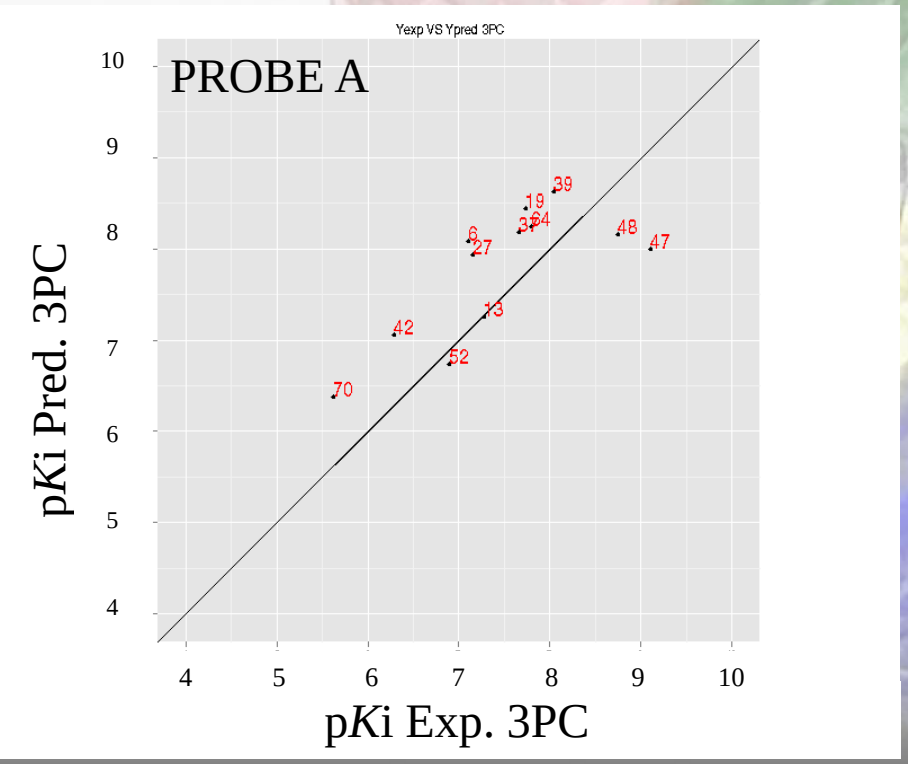
Mappa CoMFA sterica

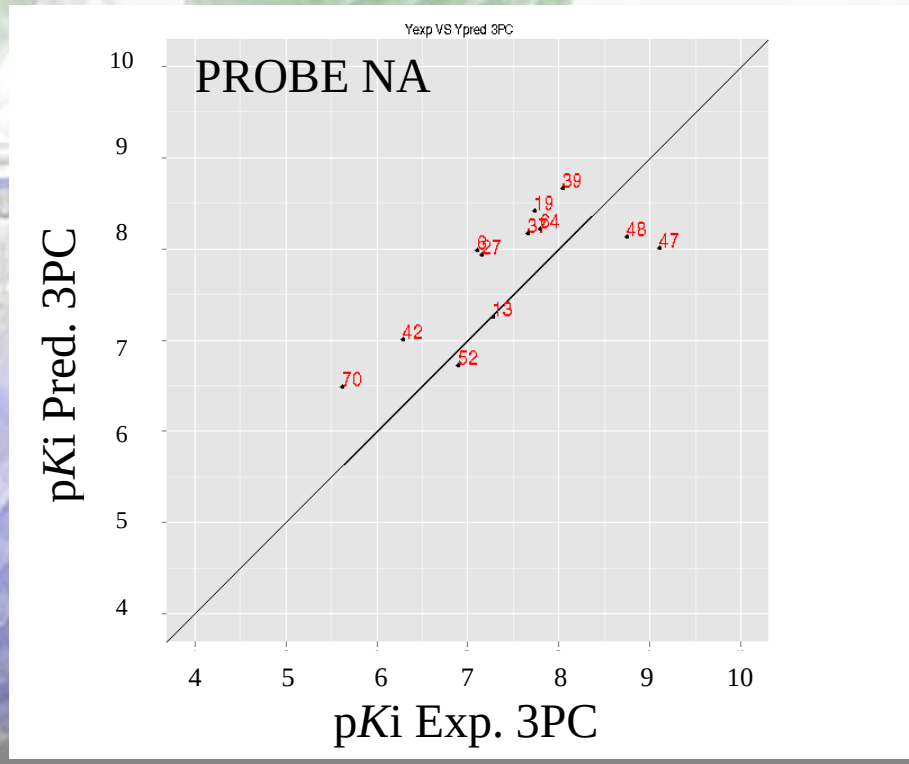


PLS-Coeff. Plot

Predizione Test Esterno 1

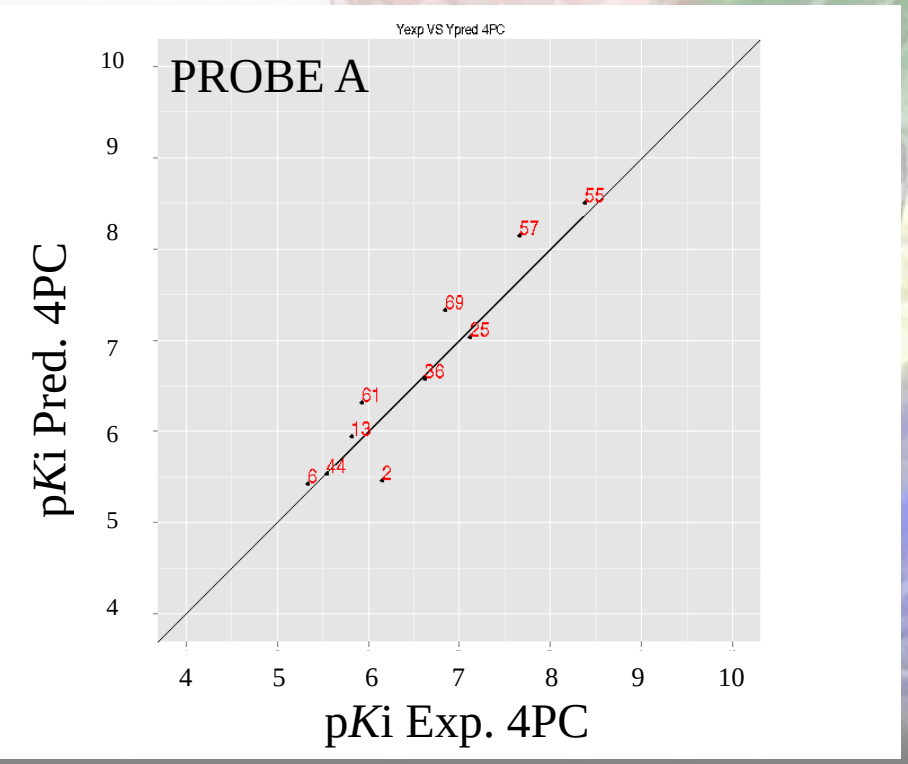
RECETTORE	N° compd TRAINING SET	N° compd TEST SET
δ	61	12
μ	48	10
κ	49	10

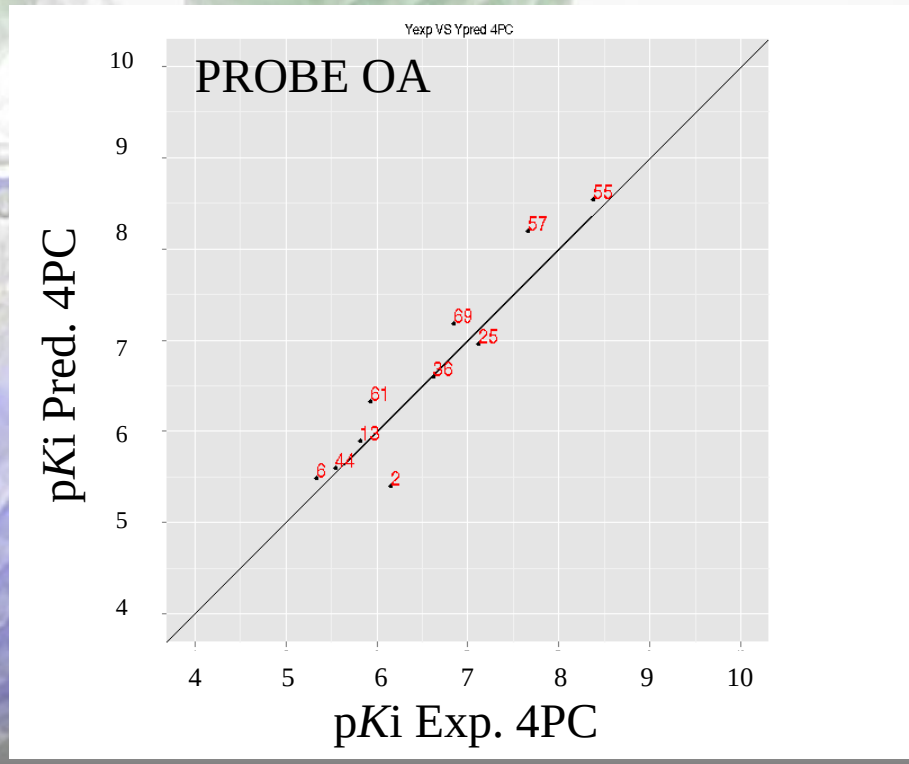




Predizione Test Esterno 1

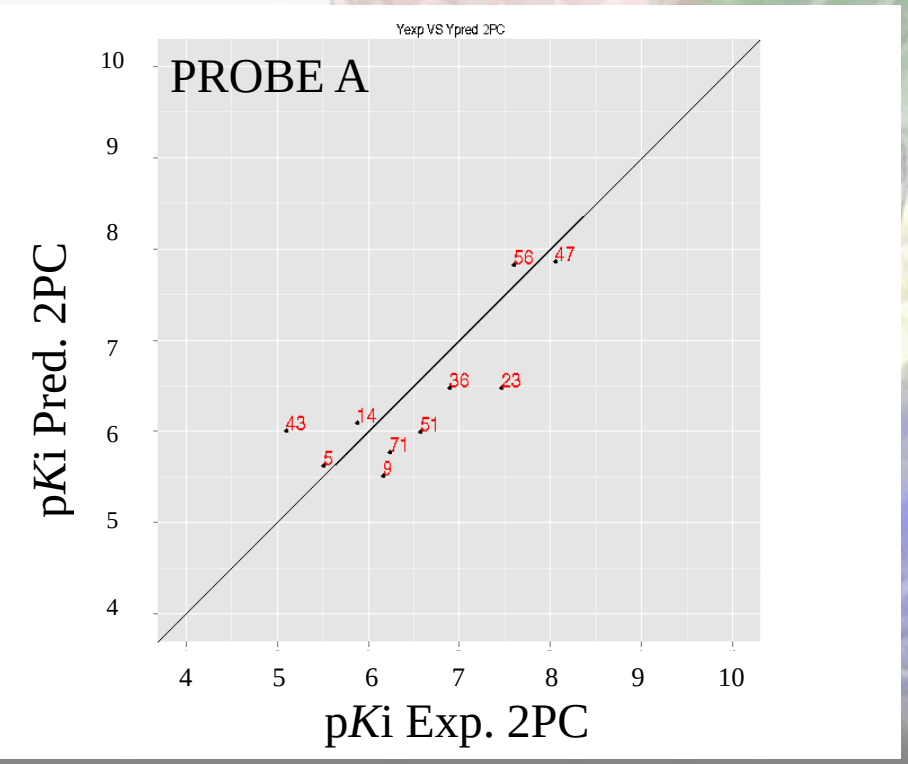
RECETTORE	N° compd TRAINING SET	N° compd TEST SET
δ	61	12
μ	48	10
κ	49	10

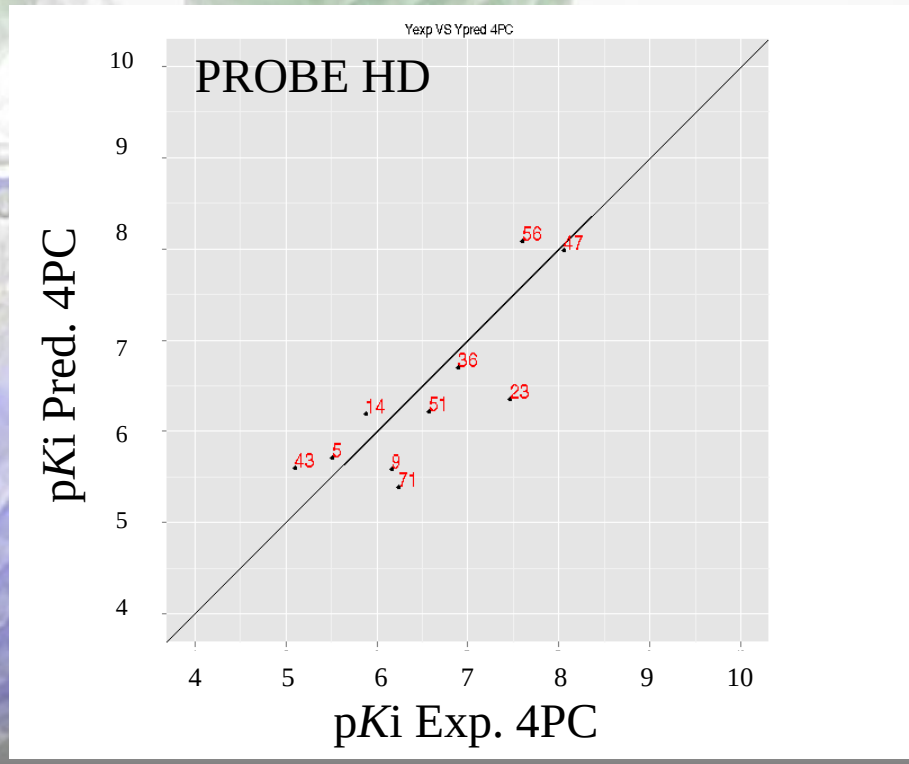




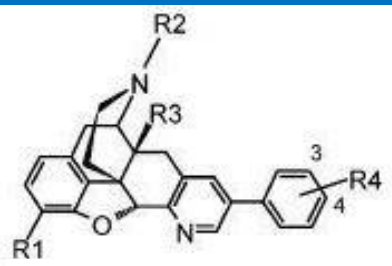
Predizione Test Esterno 1

RECETTORE	N° compd TRAINING SET	N° compd TEST SET
δ	61	12
μ	48	10
κ	49	10

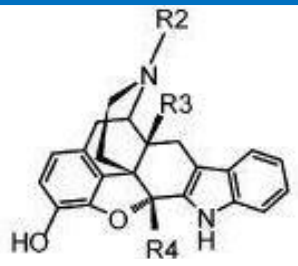




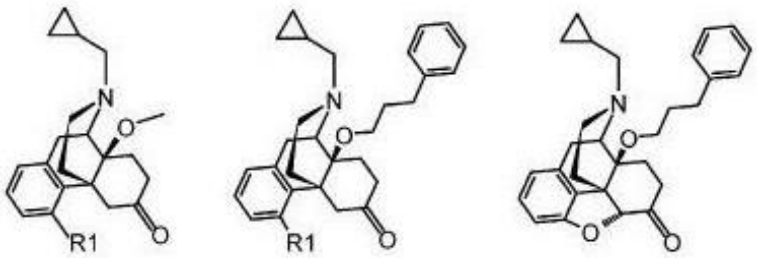
Predizione Test Set Esterno 2



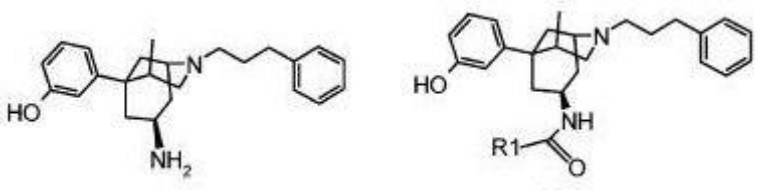
Nucleo Piridomorfinanico



Nucleo Naltreindolico



Nucleo Fenilmorfanico



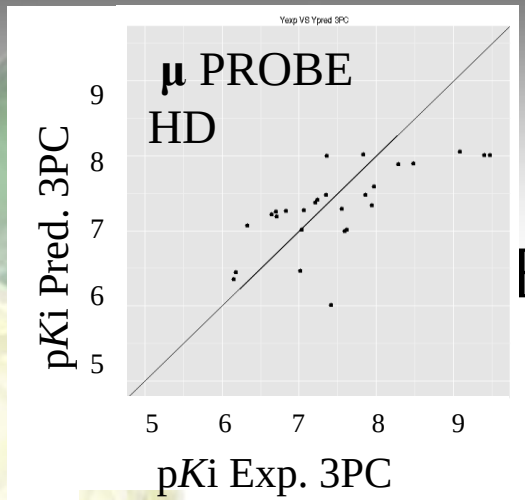
Nucleo Ciprodimico

					exptl (p*K*i)		
compd	R1	R2	R3	R4	δ	μ	κ
73	OH	allyl	OH	4-Cl	8.09	6.33	7.12
74	OH	Me	OH	H	8.54	7.59	6.44
75	OH	Me	OH	4-Cl	8.41	6.64	6.33
76	OH	Me	OH	4-Br	8.40	6.71	6.36
77	OH	Me	H	H	8.31	7.62	7.09
78	OH	Me	H	4-Cl	8.36	6.83	7.11
79	OH	Me	H	4-Br	8.30	6.70	7.04
80	OH	Me	H	3,4-Cl	8.43	7.03	6.56
81	OH	Me	H	2,4-Cl	8.96	7.01	6.39
82	OH	CPM	H	4-Cl	8.59	7.21	8.22
83	-	CPM	OEt	Me	9.10	7.41	7.23
84	-	allyl	OEt	Me	7.97	6.18	6.12
85	-	Me	OEt	Me	8.24	6.15	6.54
87	OMe				6.38	7.97	6.96
88	O(n-Bu)				6.12	7.86	6.58
89	O(cinnamyl)				6.08	7.55	6.44
90	$O(CH_2)_3Ph$				5.85	7.36	6.68
91	OCH_2Ph				6.67	7.94	7.25
92	O(n-hexyl)				5.93	7.35	6.36
93	OH				8.30	9.40	8.23
94	OMe				7.77	9.47	8.13
95	O(n-Bu)				6.63	7.83	6.82
96	-	-	-	-	7.34	9.08	7.64
97	-	-	-	-	7.83	8.48	6.91
98	$(CH_2)_3N(CH_3)_2$				5.84	7.24	7.92
99	$(CH_2)_2NHC(=NH)NH_2$				7.25	8.28	7.63
100	$(CH_2)_3NHC(=NH)NH_2$				5.76	7.06	7.88

Predizione Test Set Esterno 2

3 TIPOLOGIE DI ALLINEAMENTO:

- *Surflex*
- *Rocs*
- *Atom by Atom*



E

RECETTORE	PROBE	SDEP				
		1PC	2PC	3PC	4PC	5PC
δ	HD	**0.73**	0.74	0.85	0.74	0.73
μ	HD	1.34	1.02	**0.66**	0.70	0.66
κ	HD	0.77	**0.65**	0.71	0.65	0.73

Analisi Effettuate e dati ottenuti

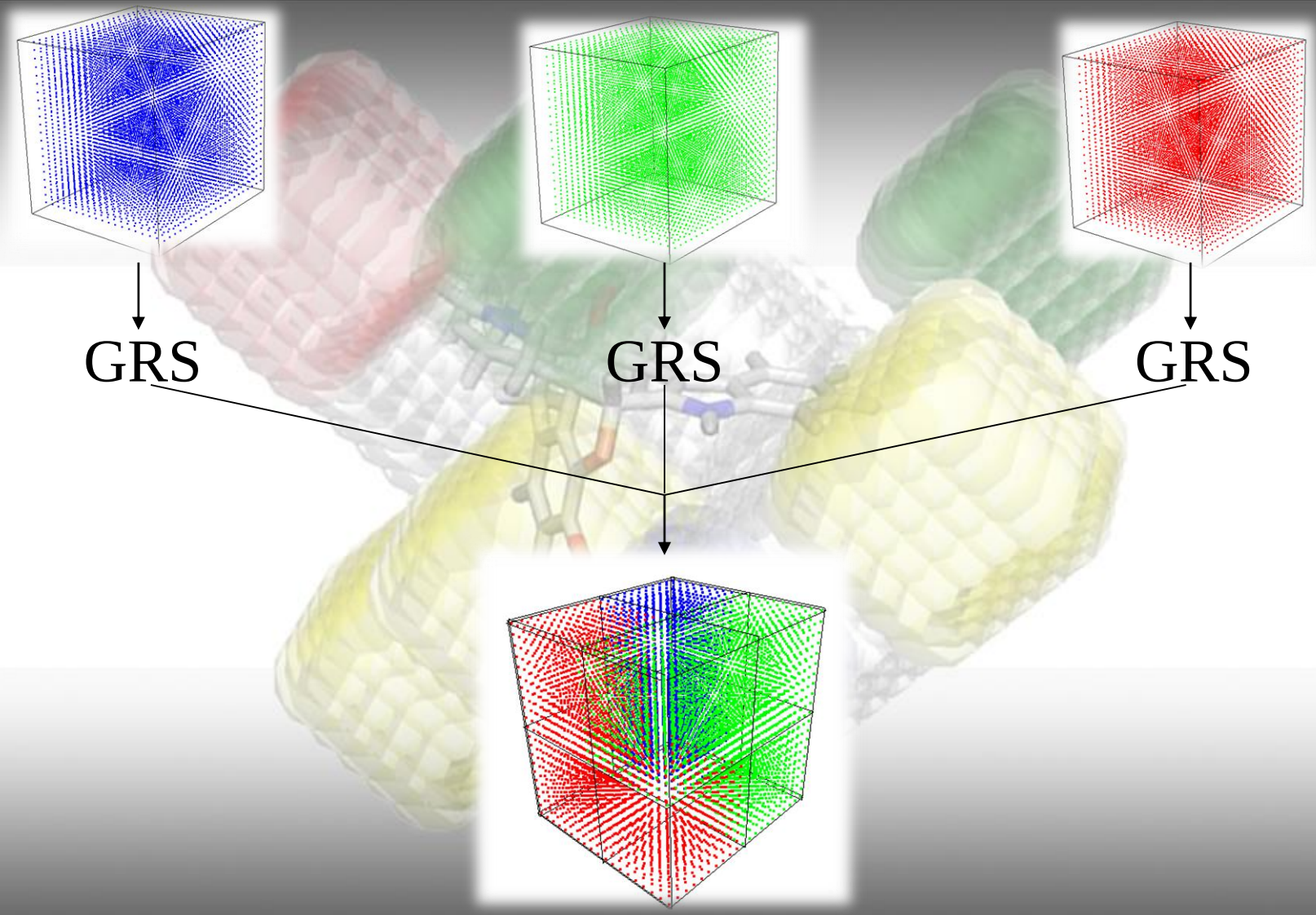
- 24 MODELLI PLS “RAW” A 5 PC
- 35056 PROCESSI DI CROSS-VALIDAZIONE A 5PC
- 293304 COMBINAZIONI DI PRET. ANALIZZATE A 5PC
- 24 MODELLI PLS “PRETRATTATI” A 5PC
- 3000 SOTTOAREE ANALIZZATE CON GRS A 5 PC
- 96 ANALISI DI PREDIZIONE

✓360 DIFFERENTI INFORMAZIONI STATISTICHE
✓2624 MAPPE DI CORRELAZIONE
✓936 PLOT 2D

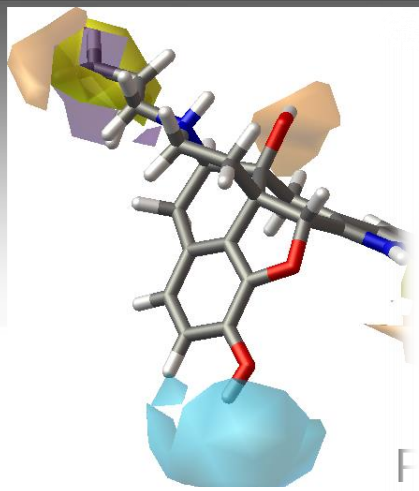
Tempo totale:
15 giorni di calcolo

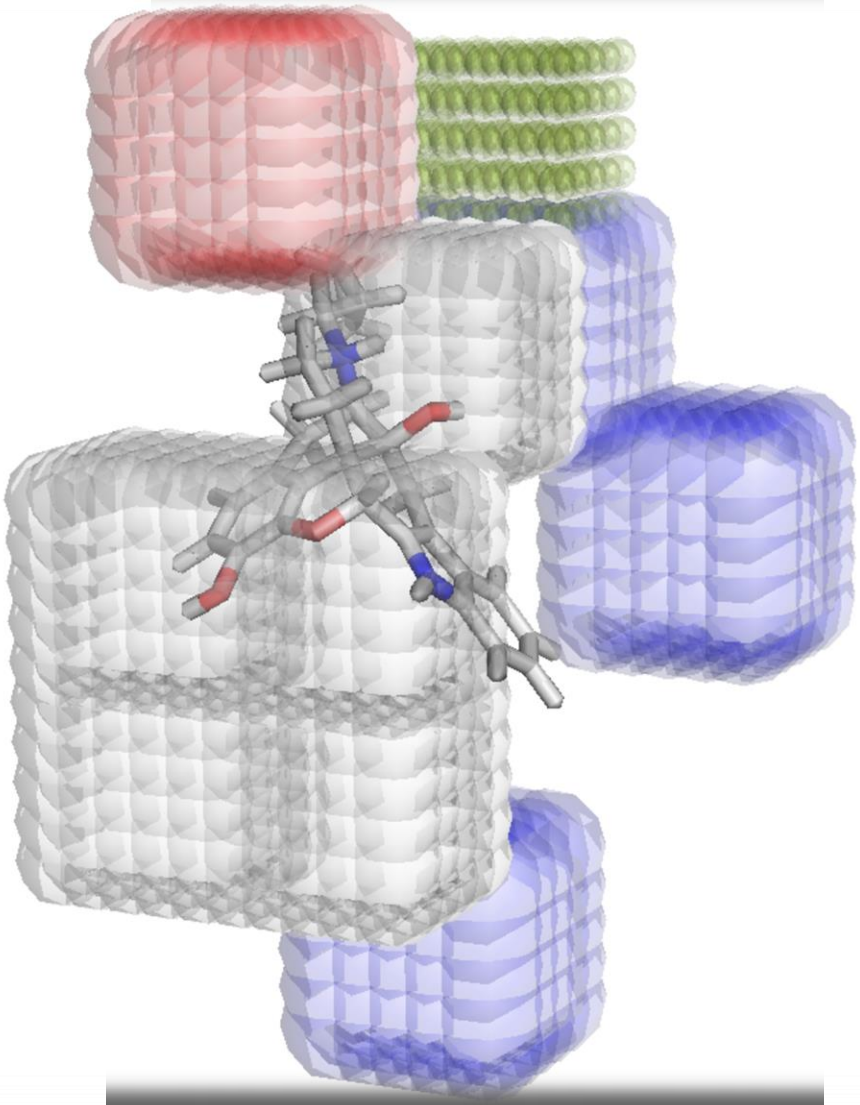
MPGRS

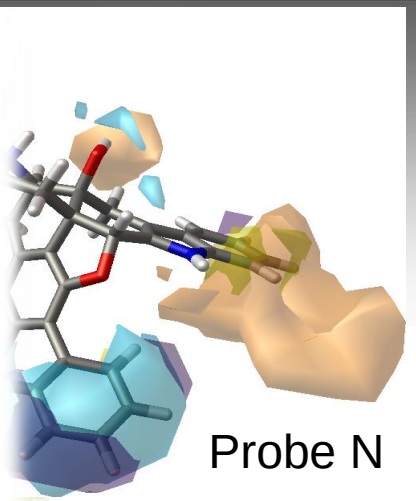
“Multi Probe Guided Region Selection”



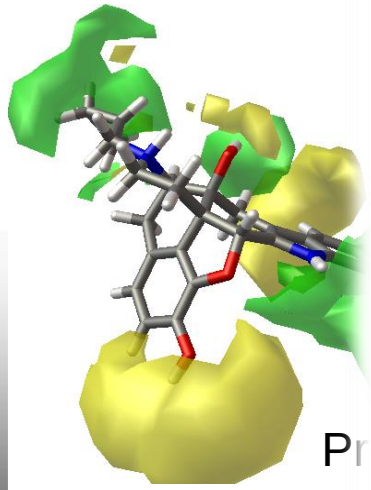
Mappe di correlazione recettore δ

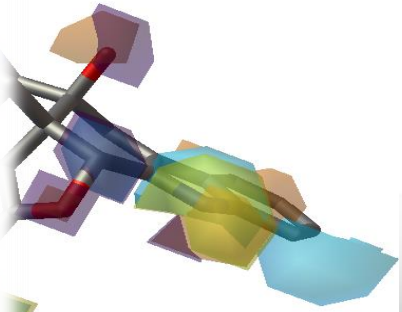






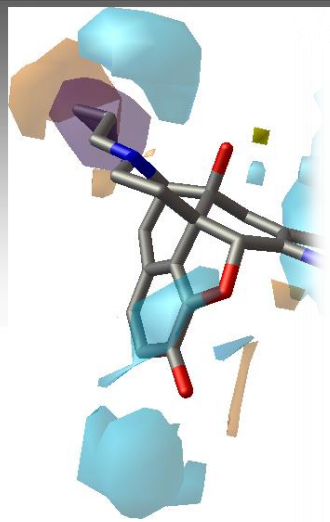
Probe N

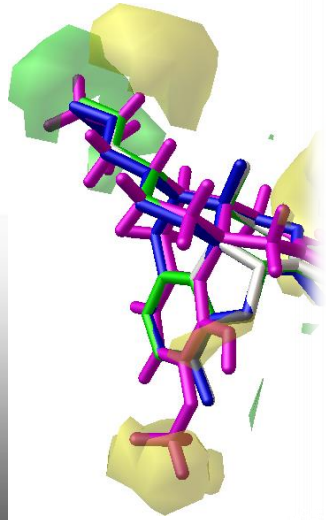


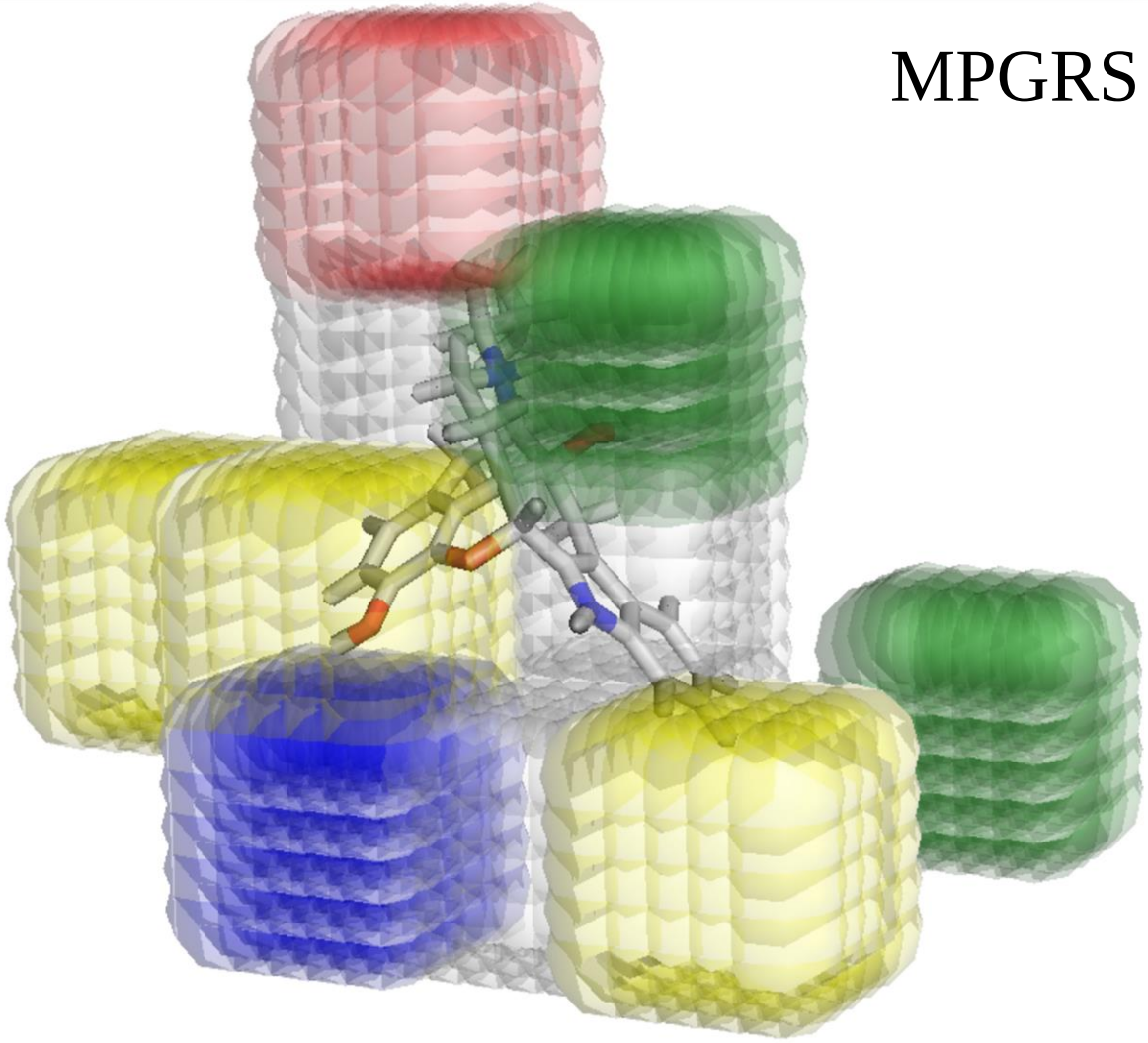


Probe e

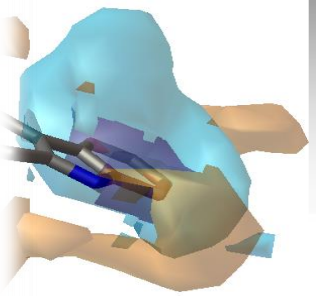
Mappe di correlazione recettore μ





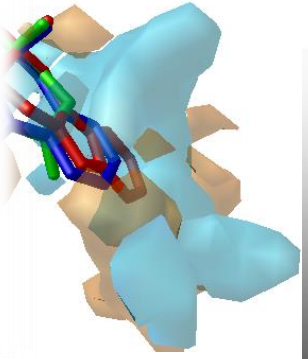


MPGRS

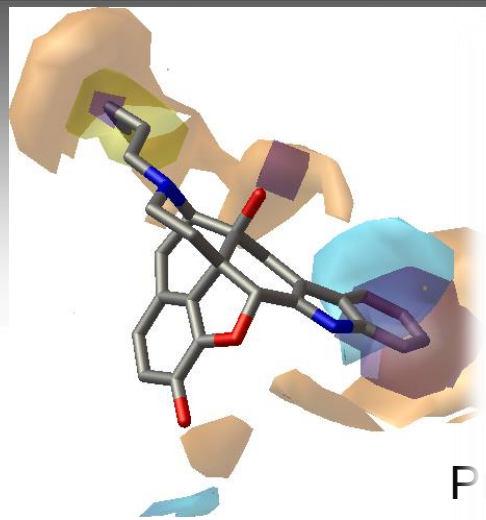


Probe C

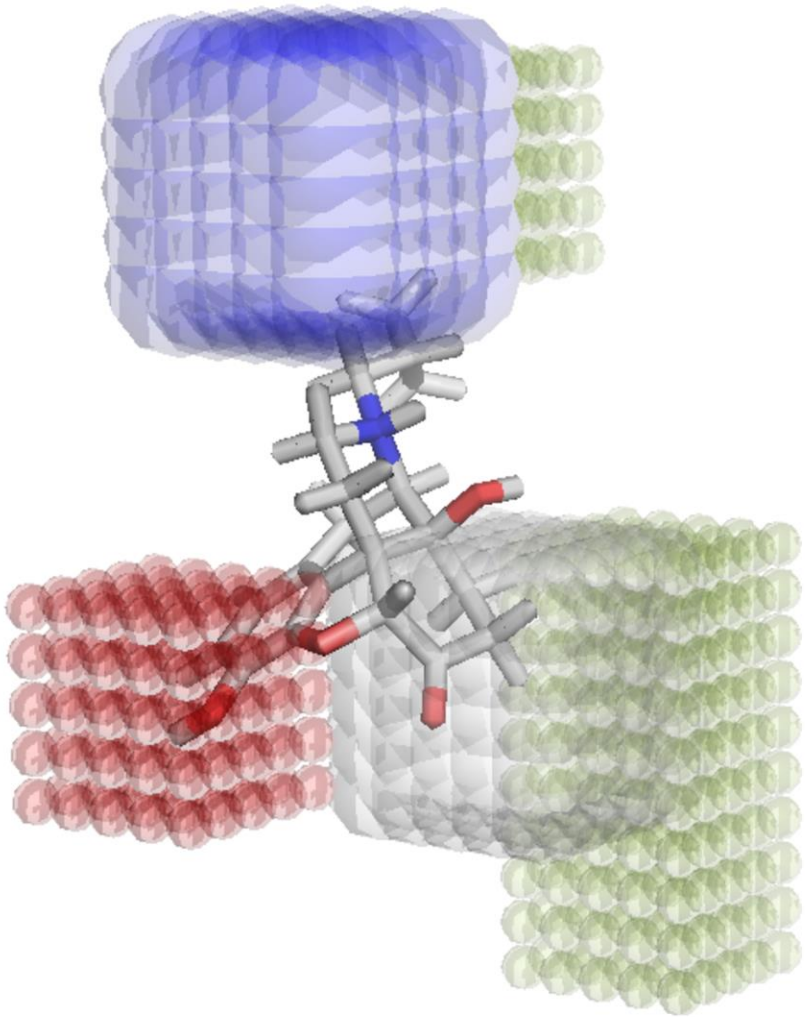
Probe HD

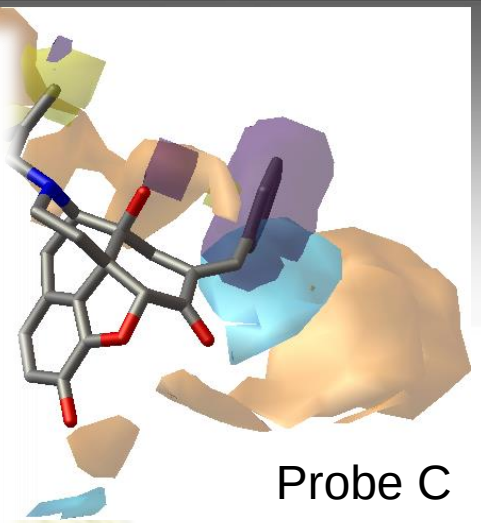


Mappe di correlazione recettore κ

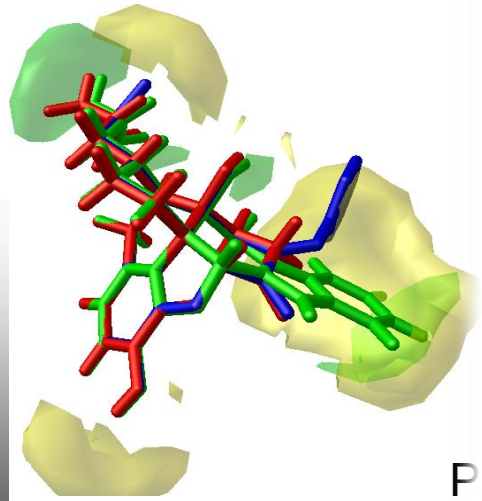


P

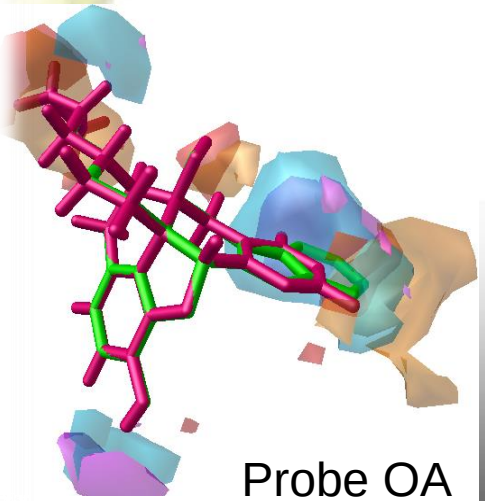




Probe C



P



Probe OA

Comunità Scientifica

"3-D QSAR SERVER – A 3-D QSAR Models Database for Virtual Screening"Alexandros Patsilinakos, Flavio Ballante, Ira Musmuca, Rino Ragno *14th Hellenic Symposium on Medicinal Chemistry April 23-25, 2010*

Flavio Ballante, Ira Musmuca, Alexandros Patsilinakos, Rino Ragno, An Alternative Method for Generating 3-D QSAR Models using Free Software. In *5th Joint Sheffield Conference on Chemoinformatics Sheffield, UK, 2010.*

"R/AUTOGRID/ADT Combination As An Alternative To Build 3-D QSAR Models. Methodologies And Applications" Musmuca Ira, Ballante Flavio, Ragno Rino *18th European Symposium on Quantitative Structure-Activity Relationships Rhodes, Greece, September 19-24, 2010*

Ringraziamenti

Ringrazio il
Prof. Rino Ragno ed il
Gruppo di Ricerca RCMD
per il supporto ricevuto durante
il corso della mia Tesi di Laurea