Applicazione di Molecular Docking, 3-D QSAR e COMBINE per lo sviluppo di modelli predittivi nell'ambito di inibitori DOT1L

Facoltà di Farmacia e Medicina

Corso di laurea in Chimica e Tecnologia Farmaceutiche

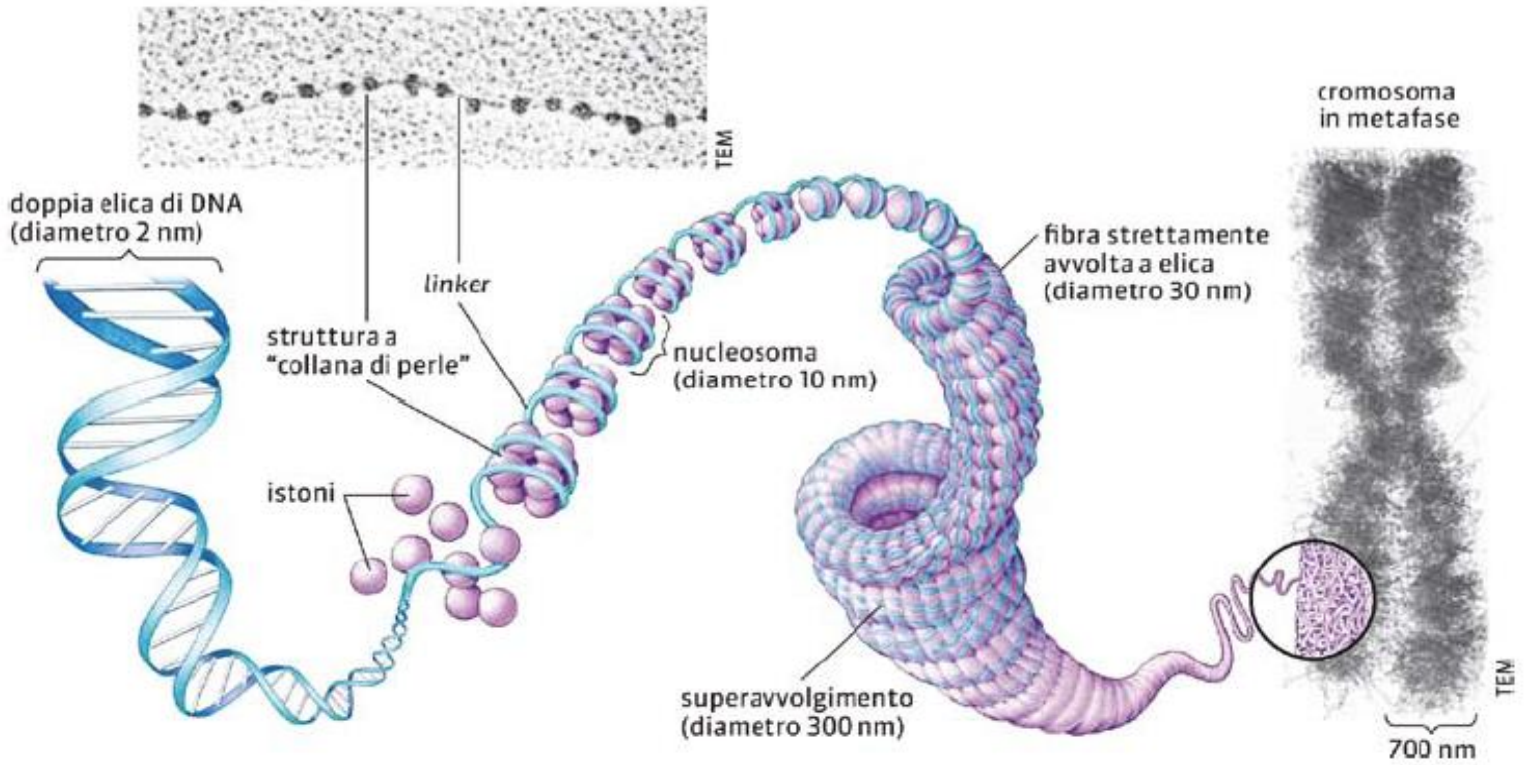
SAPIENZA
UNIVERSITÀ DI ROMA

Tesi Sperimentale in Chimica Farmaceutica
a.a. 2014/2015

Laureanda: Valentina Generoso
Matricola: 1311592

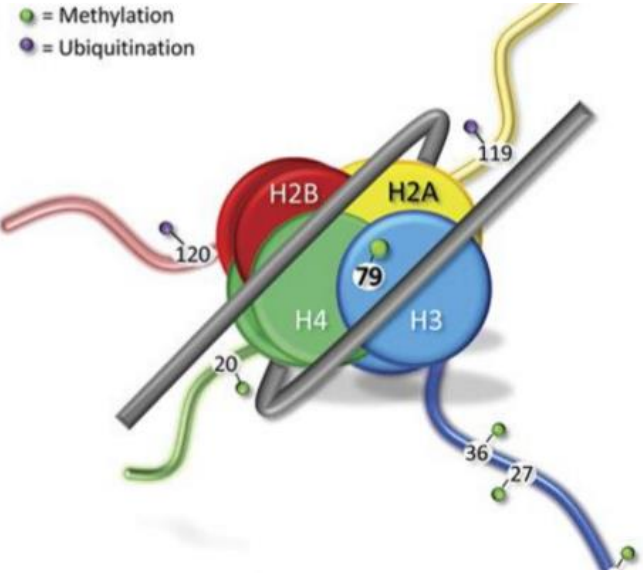
Relatore: Prof. Rino Ragno

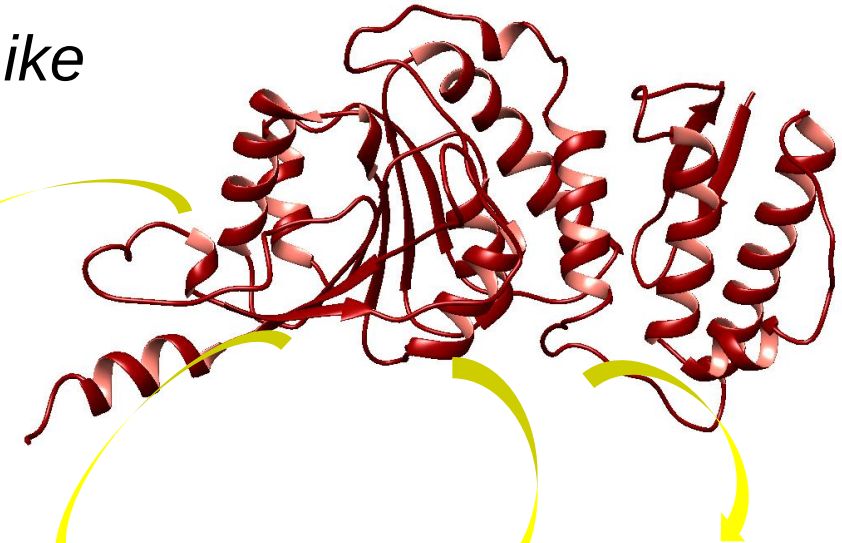
Modifiche ereditabili dell'espressione dei geni (fenotipo) che non coinvolgono cambiamenti della sequenza del DNA (genotipo)



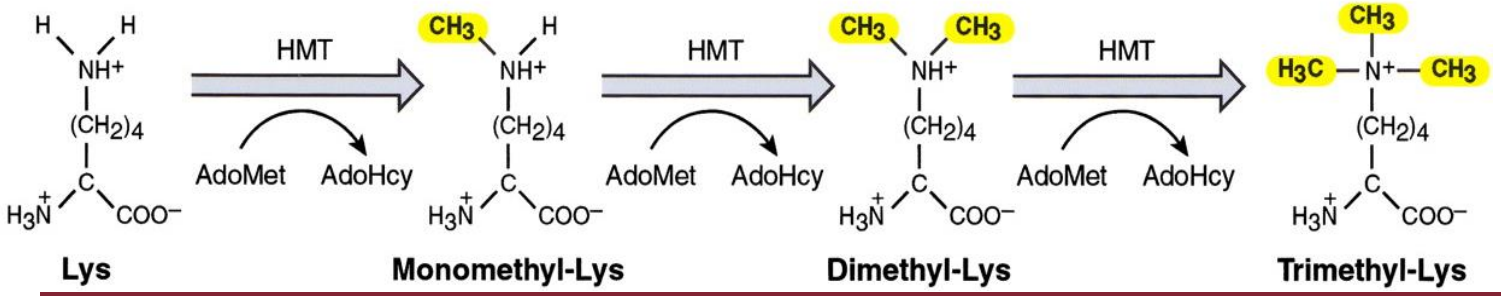
Lo stato della cromatina è regolato da modifiche specifiche a carico delle proteine istoniche e delle basi azotate. Tali modifiche costituiscono dei *marker* che vengono riconosciuti da altri complessi proteici

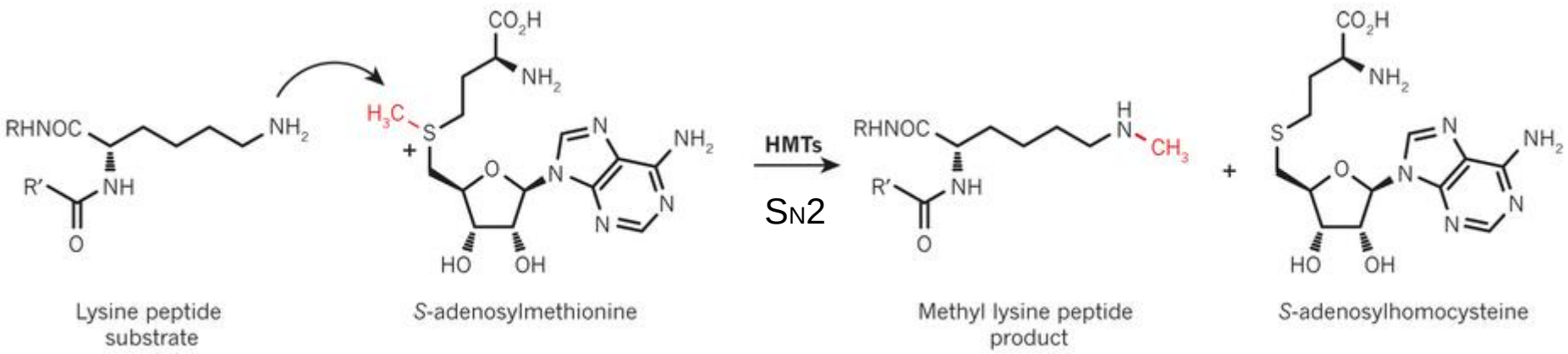
Disruptor Of Telomeric Silencing 1-Like
Metiltransferasi selettiva per H3K79



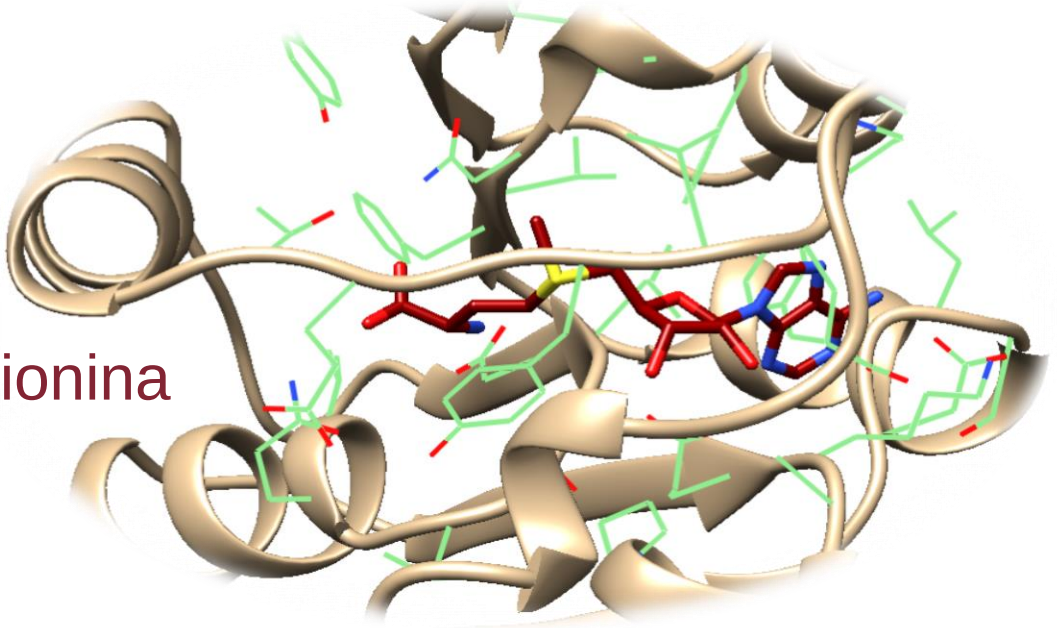


- Transizione G1-S nel ciclo cellulare
- Riparazione danni DSB
- Trascrizione
- Silenziamento telomerico



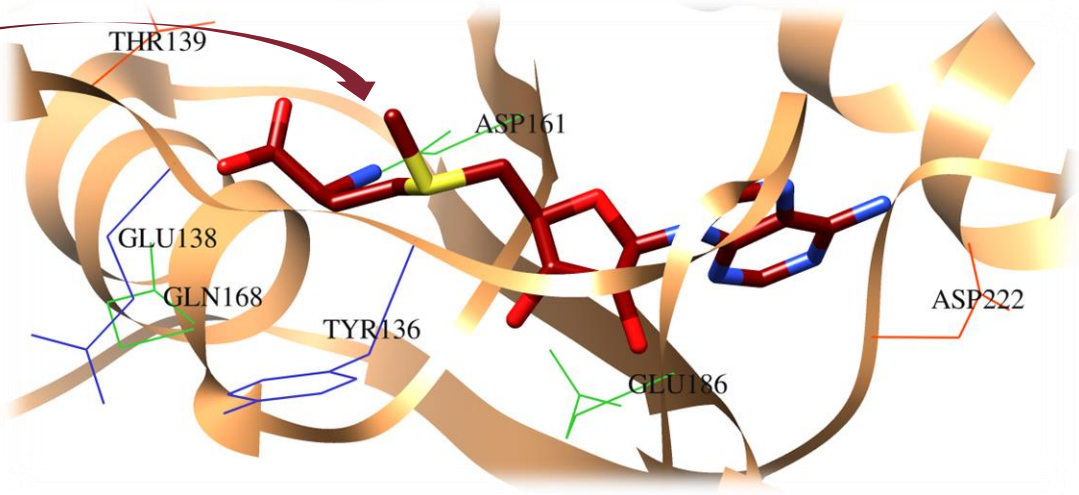


Lysine peptide substrate | *S*-adenosylmethionine | Methyl lysine peptide product | *S*-adenosylhomocysteine



Cofattore S-adenosil-L-metionina

SAM



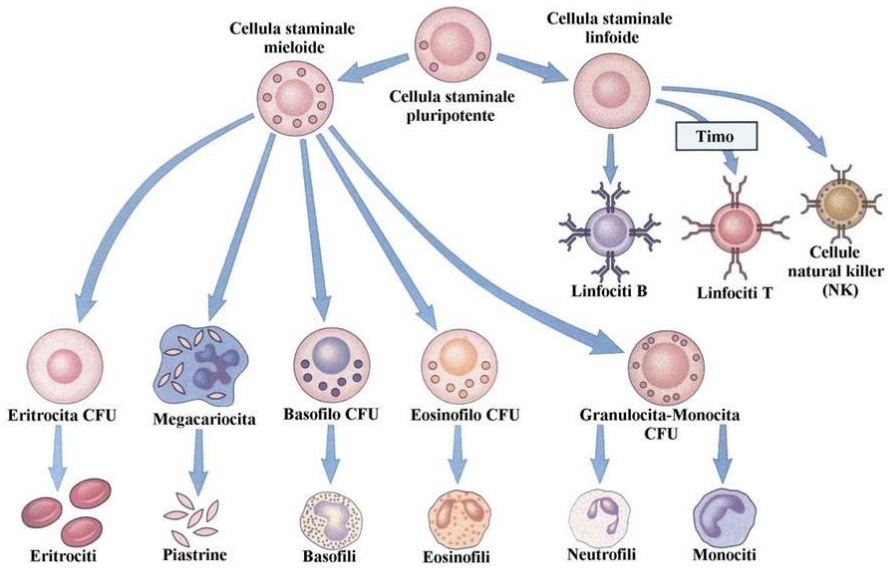
1537 aa. Sito attivo: 1-416

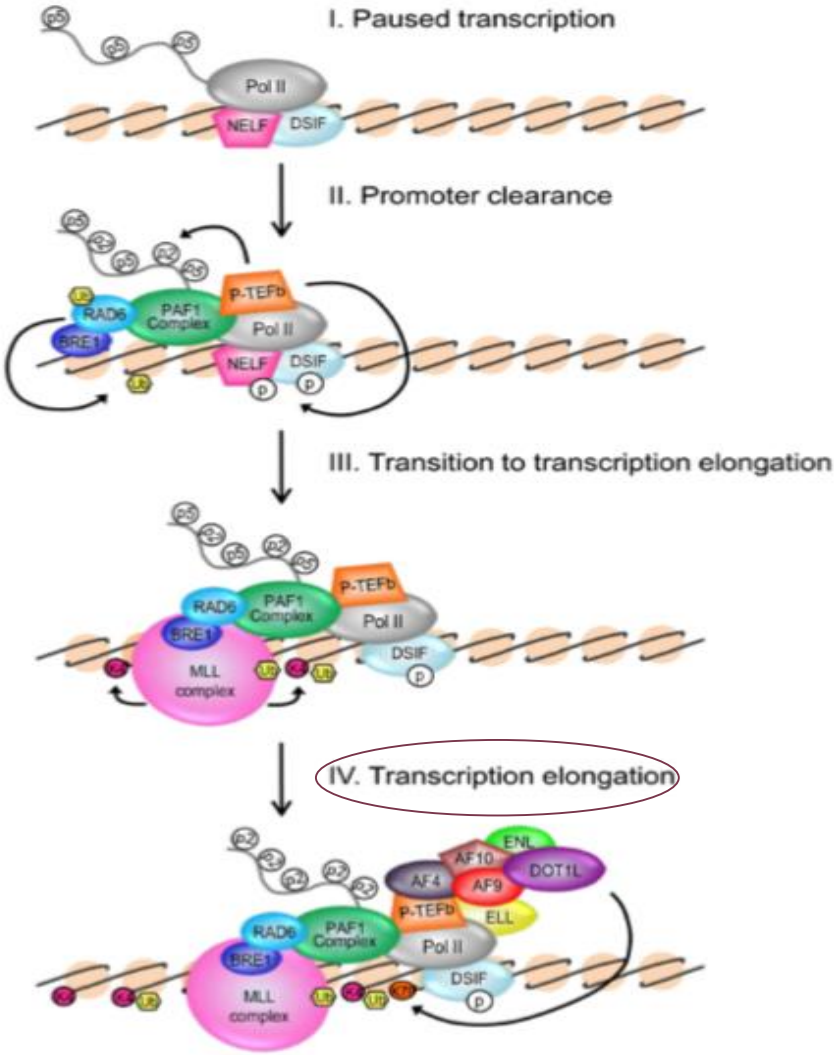
Dominio C-terminale: Struttura aperta α/β, foglietto β a 7 filamenti e 5 α eliche

Dominio N-terminale: 5 alfa eliche (αA–αF) e 2 coppie di foglietti β

SAM è alloggiata in una tasca formata da un loop che collega i due domini e dai foglietti beta del dominio C-terminale

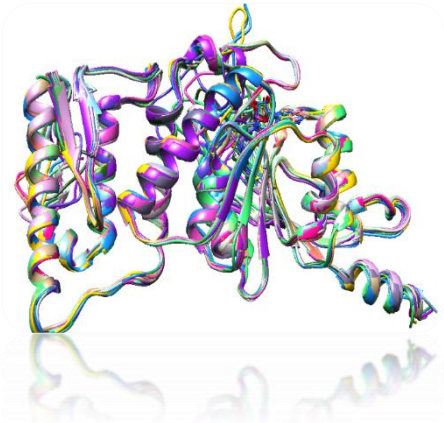
Mieloid
Linfoid
Leukemia





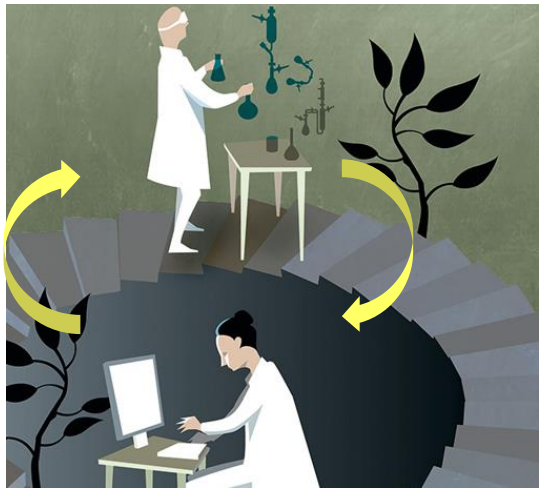
STUDIO INIBITORI DOT1L

RCSB PDB PROTEIN DATA BANK



MODELLI PREDITTIVI

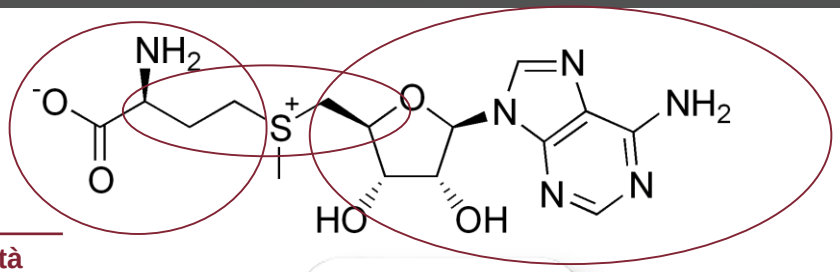
NUOVE MOLECOLE



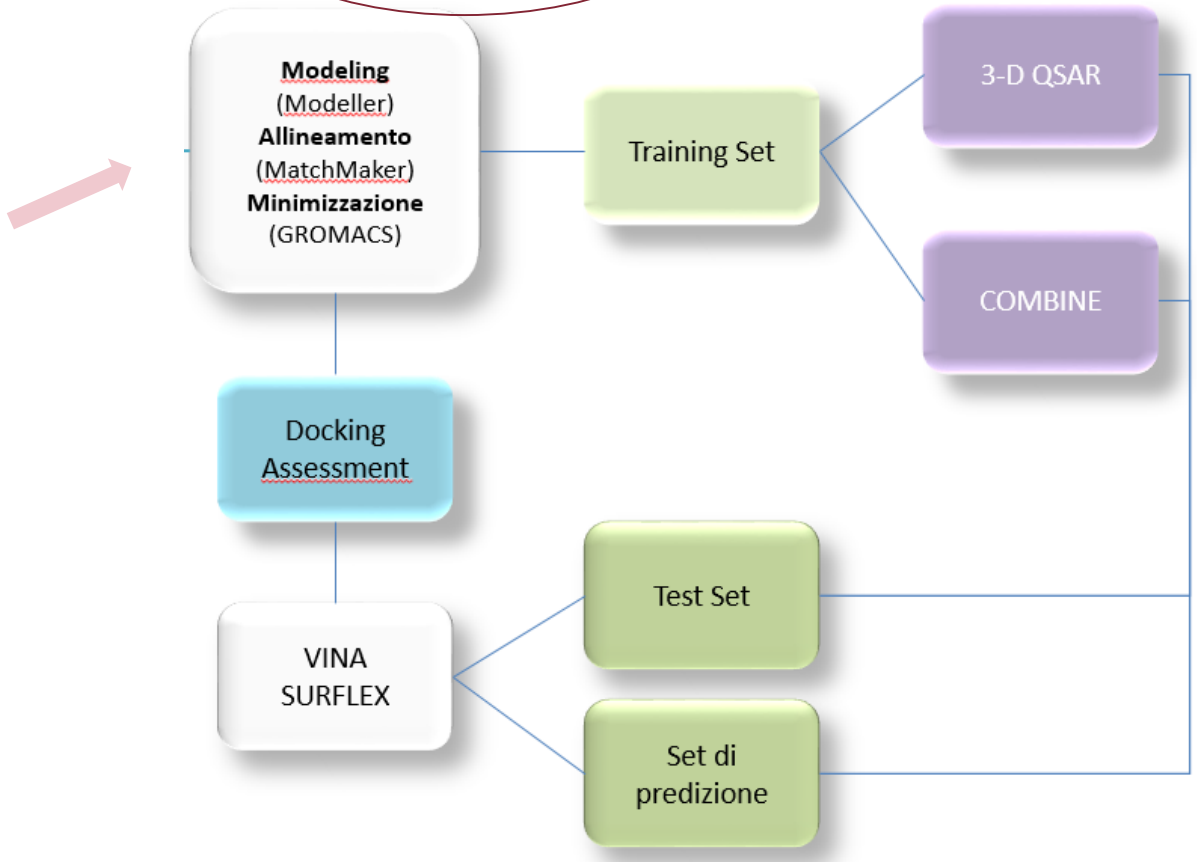
SINTESI

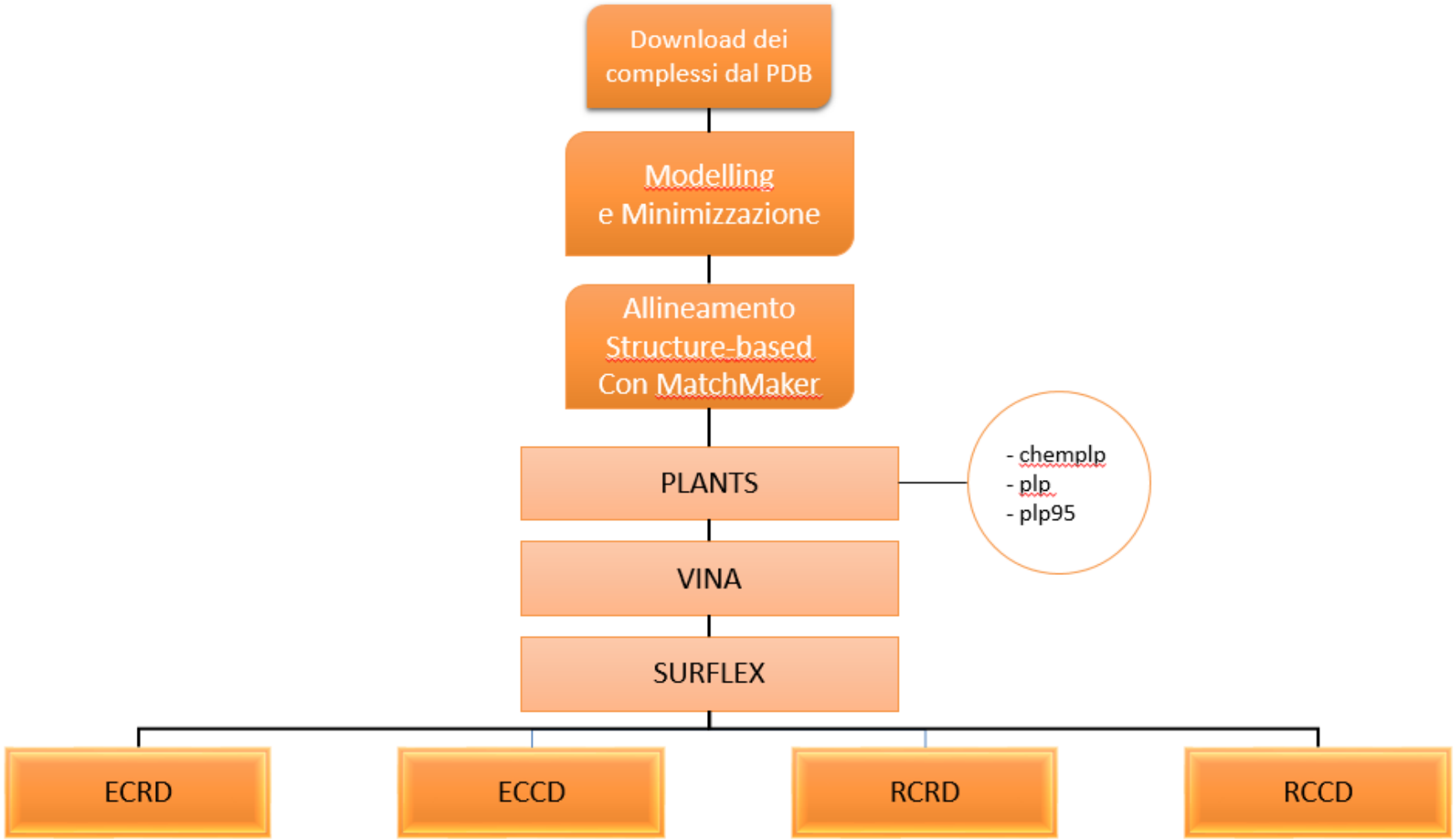
PROCEDURA SPERIMENTALE

by www.RCMD.it



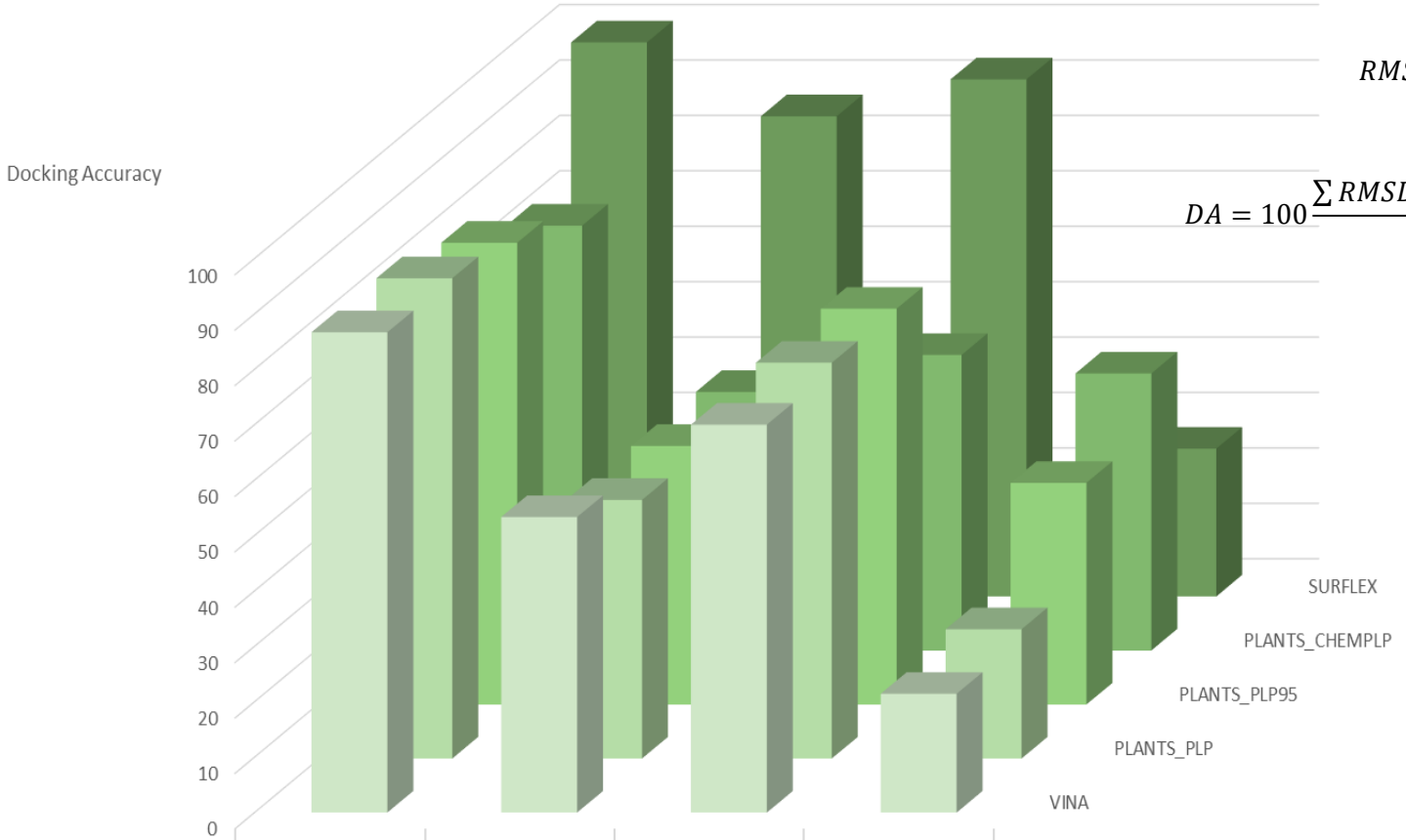
Codice ID	Codice ligando INH	Attività (pIC50 o pKi)
1NW3	SAM	6,35
3QOW	SAM	6,35
3QOX	SAH	6,57
3SR4	TT8	4,42
3SX0	SX0	7,11
3UWP	5ID	4,74
4EK9	EP4	4,42
4EKG	0QJ	7,89
4EKI	0QK	9,30
4ER5	0QK	9,30
4EQZ	AW0	7,39
4ER0	AW1	8,54
4ER6	AW2	9,52
4ER7	AW3	8,44
4HRA	EP6	10,10




Download dei complessi dal PDB
Modelling e Minimizzazione
Allineamento Structure-based Con MatchMaker
PLANTS
- chemplp
- plp
- plp95
VINA
SURFLEX
ECRD
ECCD
RCRD
RCCD

DOCKING ASSESSMENT

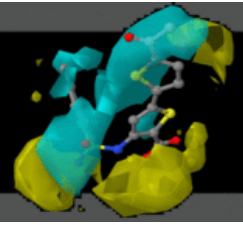
by www.RCMD.it



$$RMSD = \sqrt{\frac{\sum_{i=1}^{N} d_i^2}{N}}$$

$$DA = 100\frac{\sum RMSD_{(0-2)} + \sum RMSD_{(2-3)}}{N}$$

	ECRD	ECCD	RCRD	RCCD
VINA	86,67	53,33	70	21,43
PLANTS_PLP	86,67	46,67	71,43	23,33
PLANTS_PLP95	83,33	46,67	71,43	40
PLANTS_CHEMPLP	76,67	46,67	53,33	50
SURFLEX	100	86,67	93,33	26,67

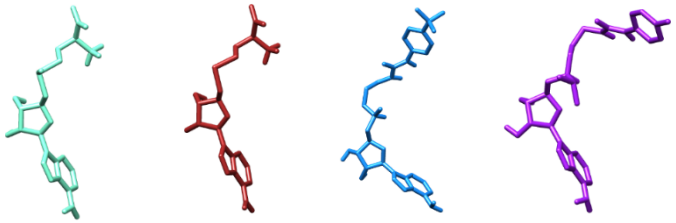


APPLICAZIONE DOCKING

by www.RCMD.it

External Test Set

3QOX 3SR4 4EKG 4EQZ



3QOX_AA 4.00
3QOX_BB 4.92

3SR4_CC 5.96
3SR4_DD 4.80
3SR4_EE 7.42
3SR4_FF 6.92
3SR4_GG 6.96

4EKG_II 4.70
4EKG_JJ 6.07
4EKG_KK 8.40

4EQZ_LL 4.34
4EQZ_MM 4.66
4EQZ_NN 4.80
4EQZ_OO 4.66
4EQZ_PP 5.72
4EQZ_QQ 6.26
4EQZ_RR 4.74
4EQZ_SS 5.96
4EQZ_TT 7.09
4EQZ_UU 7.15
4EQZ_VV 7.24
4EQZ_WW 9.10
4EQZ_XX 8.96
4EQZ_YY 8.89

Analisi DA%

Surflex

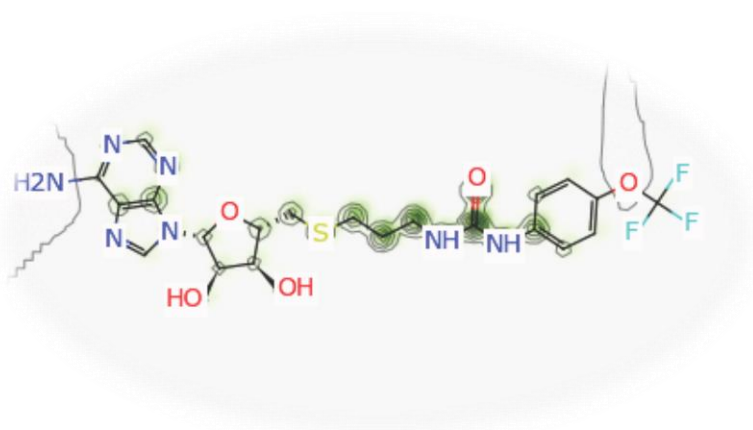
Vina

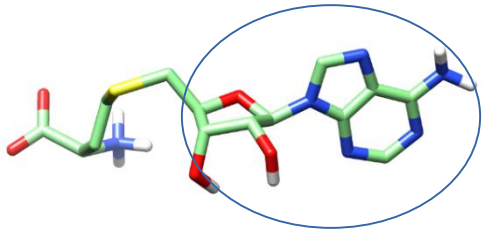
cross-docking

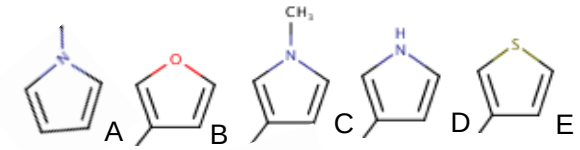
3-D QSAR

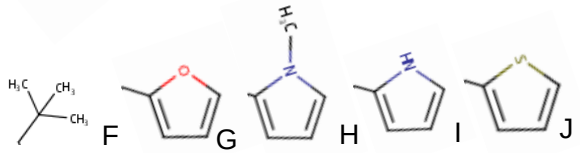
COMBINE

Prediction Test Set

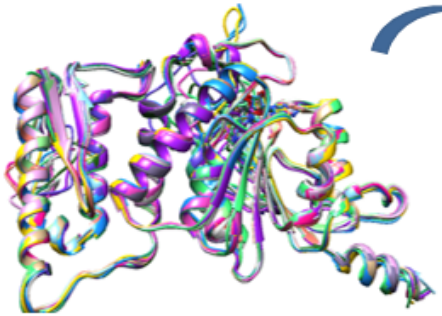




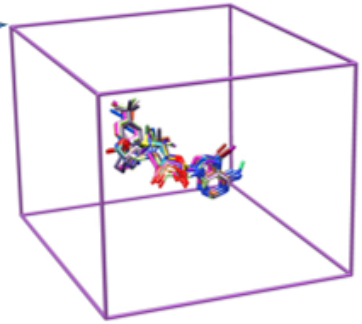




Codice ID	Codice ligando INH
1NW3	SAM
3QOW	SAM
3QOX	SAH
3SR4	TT8
3SX0	SX0
UWP	5ID
4EK9	EP4
4EKG	0QJ
4EKI	0QK
4ER5	0QK
4EQZ	AW0
4ER0	AW1
4ER6	AW2
4ER7	AW3
4HRA	EP6



Allineamento e minimizzazione



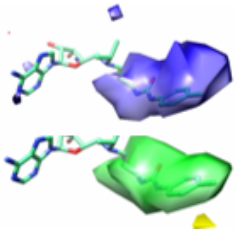
3-D QSAR metodo PLS

		FITTING		PRET		Pretreatment		
Probe	PC	r^2	SDEC	q^2	SDEP	CutOff	Zeroing	MinStd
A	1	0.797	0.817	0.653	1.069	2	0.006	0.005
C	1	0.797	0.818	0.651	1.072	2	0.01	0.005
HD	1	0.768	0.875	0.616	1.124	5	0.007	0.05
OA	4	0.989	0.187	0.728	0.947	4	0.004	0.005
NA	3	0.936	0.46	0.693	1.006	4	0.002	0.005
N	3	0.938	0.452	0.675	1.034	5	0.002	0.005
e	2	0.767	0.877	0.392	1.416	4	0.009	0.05
d	1	0.851	0.7	0.804	0.803	1	0.009	0.01

Validazione mediante External Test Set

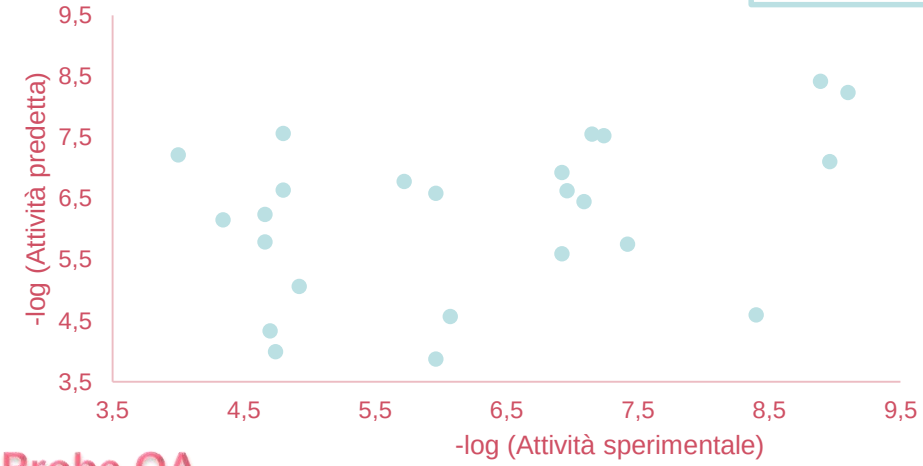
Predizione delle attività di un set di 90 molecole (Prediction Test Set)

Interpretazione Dei modelli



Attività predette (Vina)

External Test Set

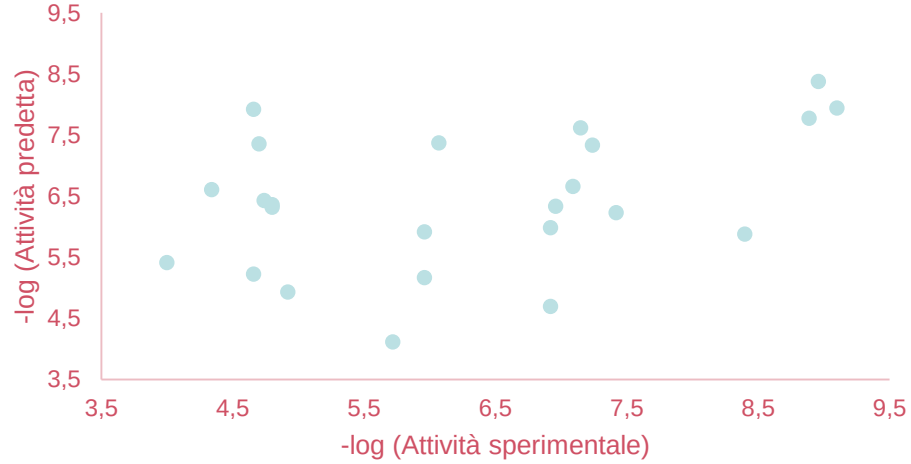


$$SDEP = \sqrt[2]{\frac{\sum(Y - Y')^2}{N}}$$

Permette di valutare la capacità predittiva del modello

Probe OA LOO

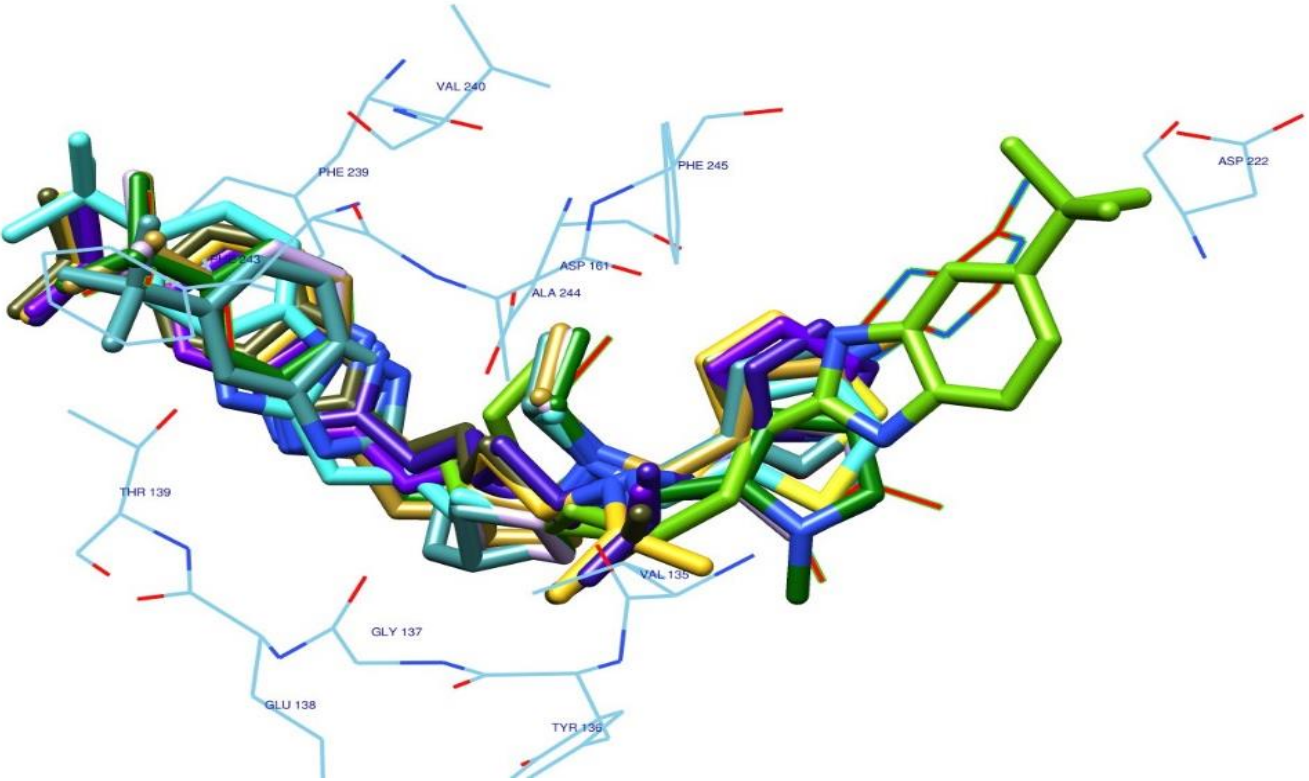
Attività predette (Surflex)



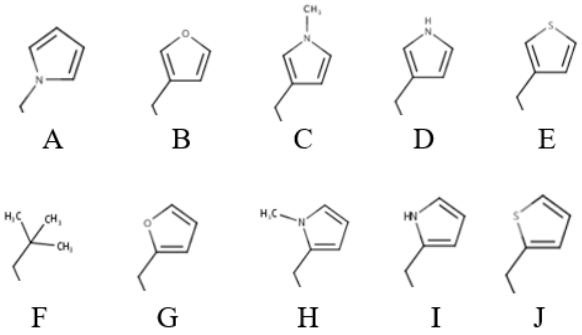
Probe	*SDEP VINA*	*SDEP SURFLEX*
A	1.58	1.76
C	1.58	1.76
HD	1.59	1.84
OA	1.60	1.54
NA	1.49	1.65
N	1.49	1.62
e	1.6	1.82
d	1.48	2.06
Media	1.55	1.76

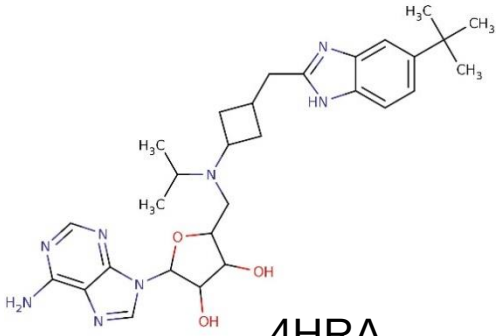
Probe OA
LOO

Prediction Test Set

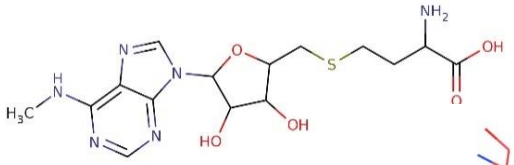


COMPOSTO	attività
4HRA	10,10
4HRA_A	7,31
4HRA_B	7,10
4HRA_C	7,07
4HRA_D	9,26
4HRA_E	8,80
4HRA_F	9,16
4HRA_G	5,53
4HRA_H	9,06
4HRA_I	7,16
4HRA_J	8,27





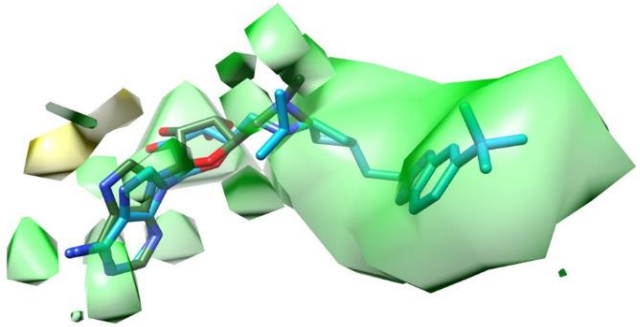
4HRA



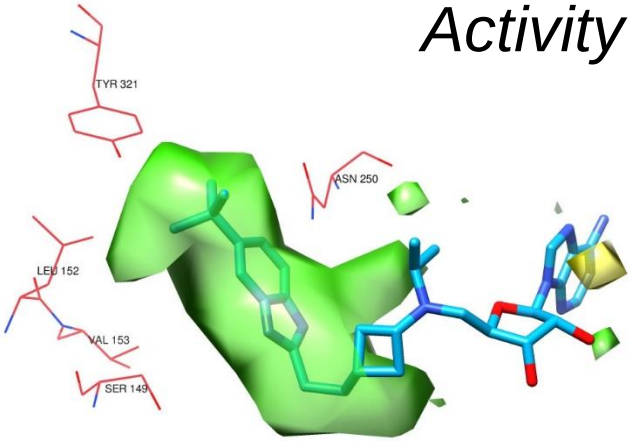
3SR4

Probe C

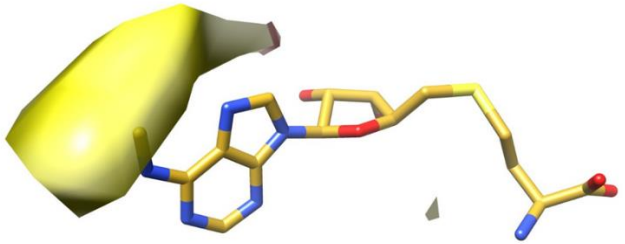
Average Activity Contribution Plot



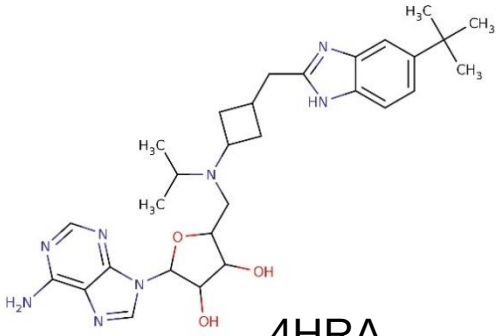
Activity Contribution Plot



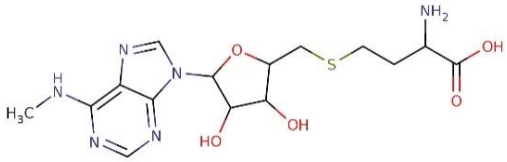
4HRA pKi = 10,10



3SR4 pKi = 4,42



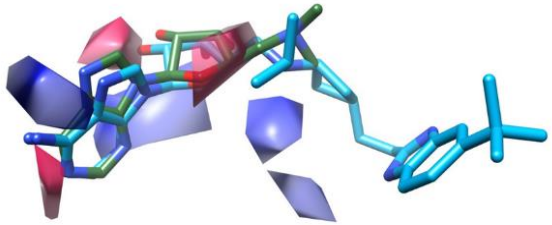
4HRA

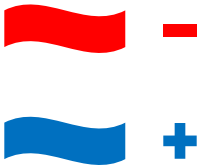


3SR4

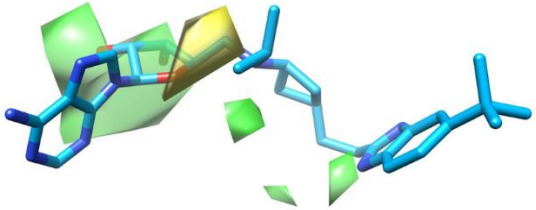
Probe e

Average Activity Contribution Plot

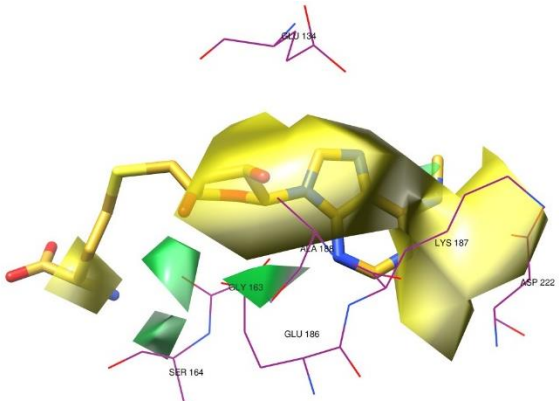




Activity Contribution Plot



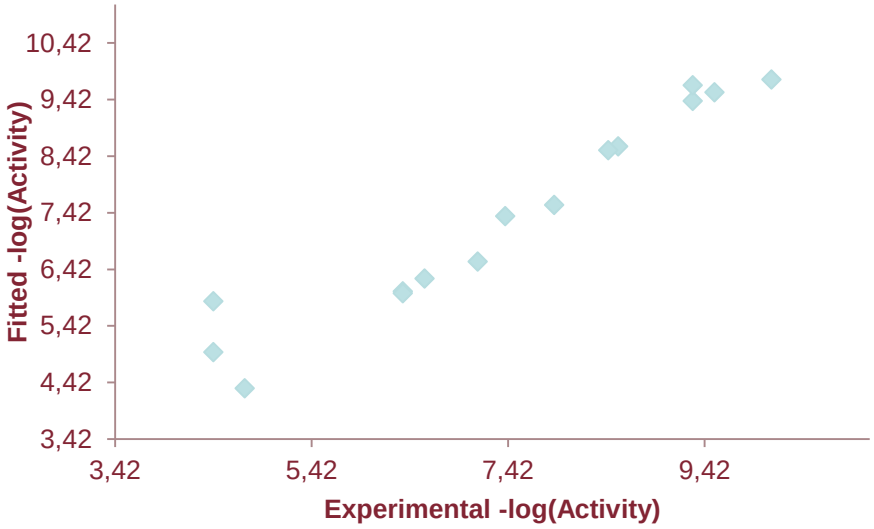
4HRA pKi = 10,10



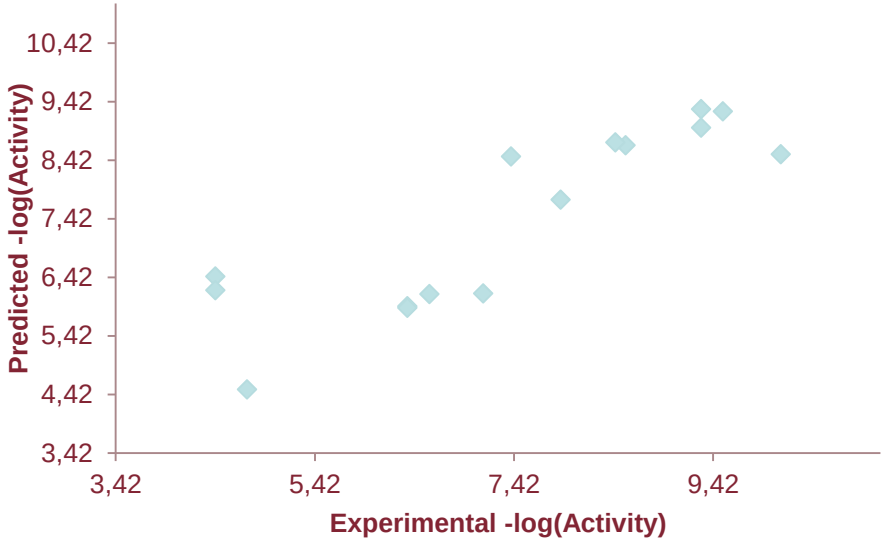
3SR4 pKi = 4,42

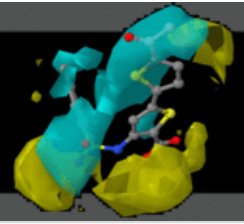
INTERAZIONI	PC	r2	q2
STE-DRY-ELE-HB	3	0,93	0,74

Fitting

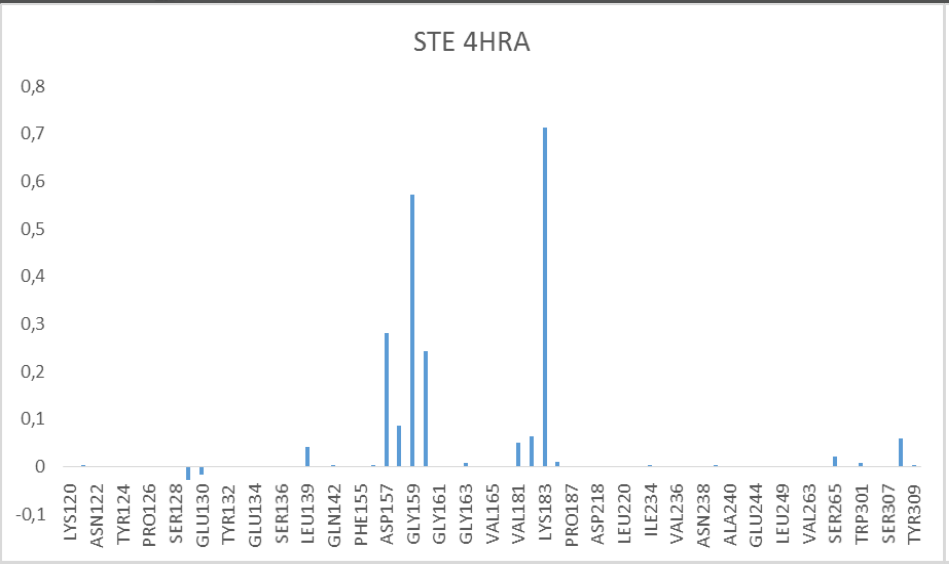


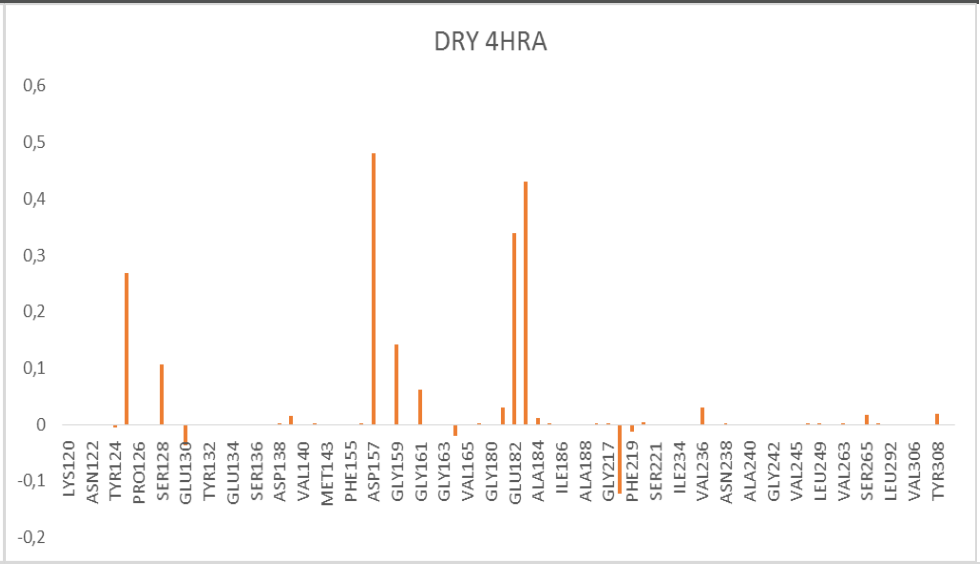
Crossvalidation

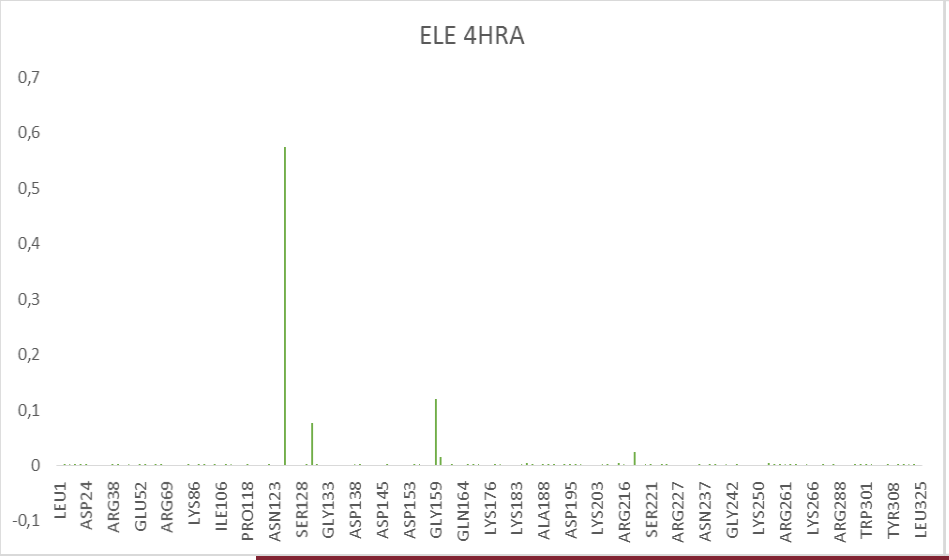


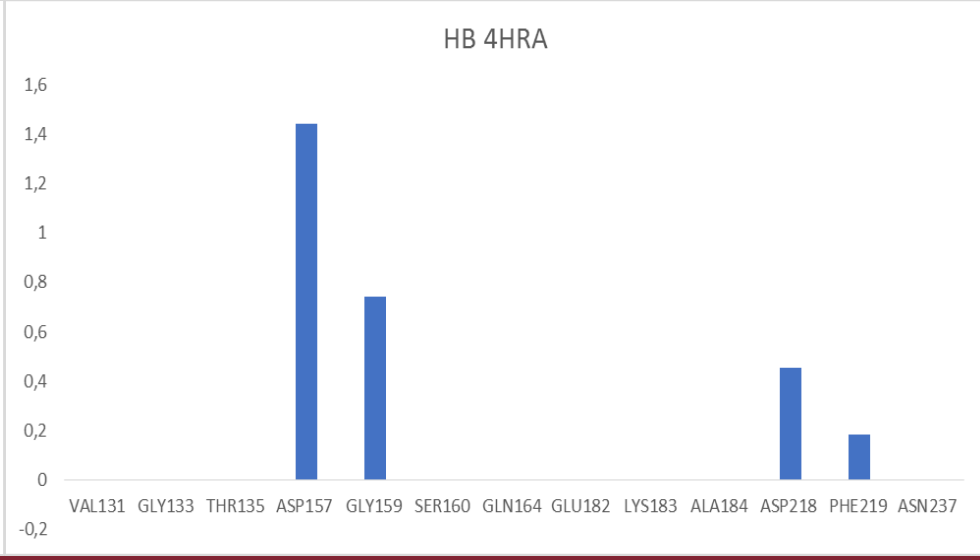


COMBINE (simulated annealing) by www.RCMD.it

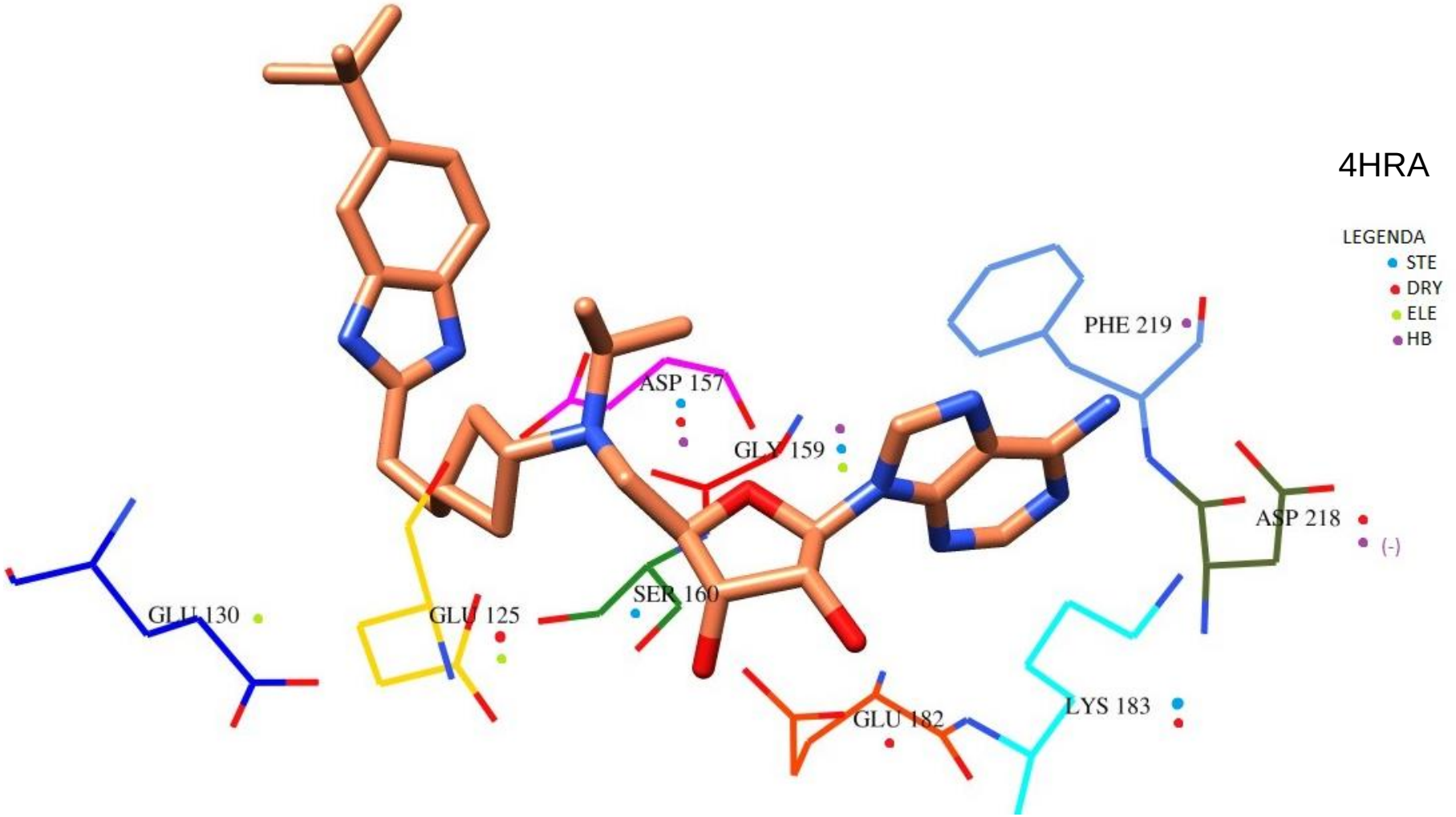


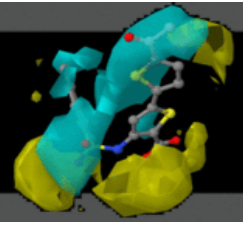






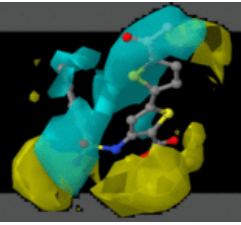
COMBINE (simulated annealing)





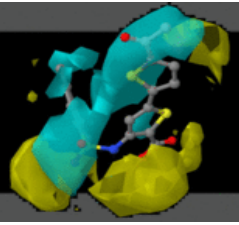
CONCLUSIONI

- Determinazione dei programmi di docking che mostrano la performance migliore: Vina e Surflex
- Sviluppo di modelli 3-D QSAR e COMBINE robusti e affidabili
- Predizione delle attività di set di molecole
- Informazioni utili per lo studio del binding mode e delle interazioni all'interno della tasca recettoriale



OBIETTIVI FUTURI

- Ampliare il training set per rendere i modelli ancora più validi
- Utilizzo dei modelli per la predizione di nuove librerie di molecole
- Progettazione di nuovi inibitori



Grazie per l'attenzione