Riproducibilità di studi ligand-based e structure-based. Applicazione ad una serie di ligandi allosterici delle TYK2.

SAPIENZA
UNIVERSITÀ DI ROMA

Facoltà di Farmacia e Medicina
Corso di Laurea in Farmacia
Tesi Sperimentale in Chimica Farmaceutica

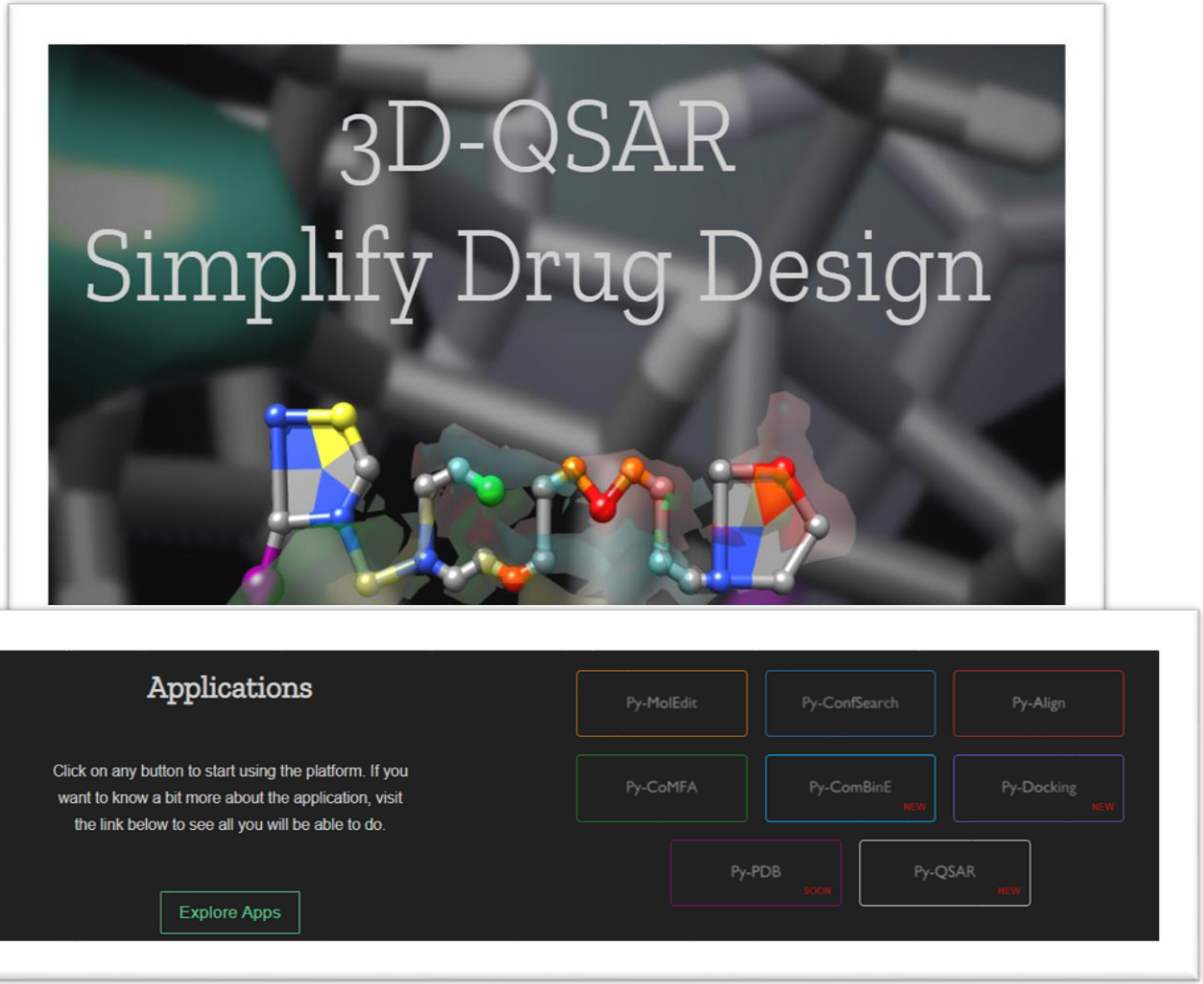
a.a 2020/2021

Laureando: Sveva Nigro
Matricola: 1534878

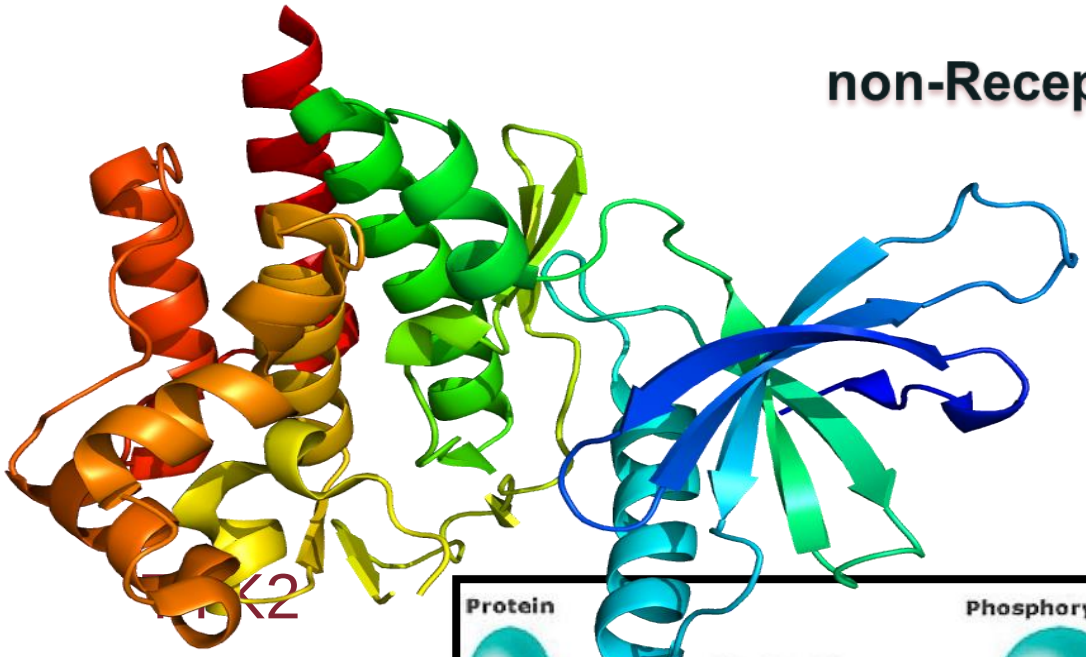
Relatore: prof. Rino Ragno

Scopo della tesi

Riproducibilità di studi ligand-based e structure-based. Applicazione ad una serie di ligandi allosterici delle TYK2.

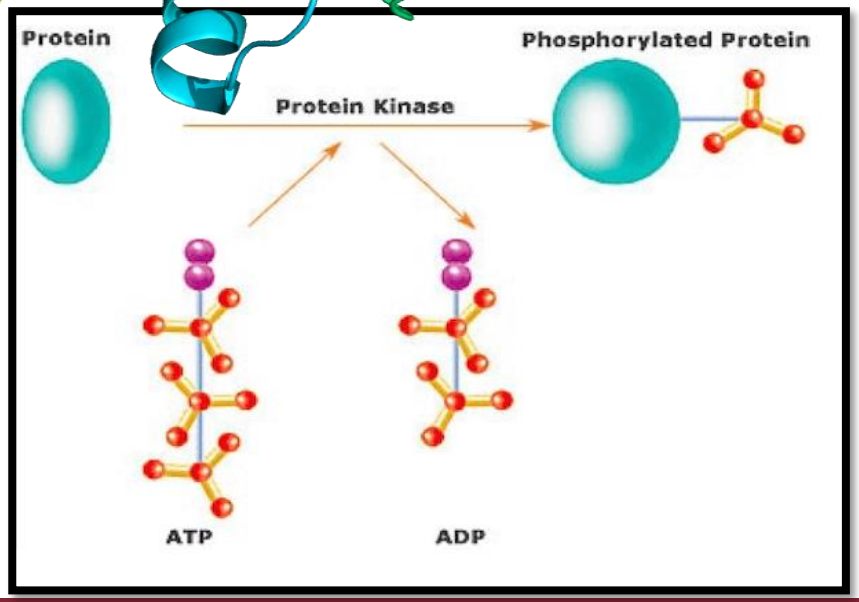


Tirosin chinasi non recettoriali



non-Receptor Tyrosine Kinases (NRTKs)

- ABL
- ACK
- CSK
- FAK
- FES
- FRK
- **JAK**
  - **JAK1**
  - **JAK2**
  - **JAK3**
  - **TYK2**
- SRC
- TEC
- SYK



Struttura delle JAK

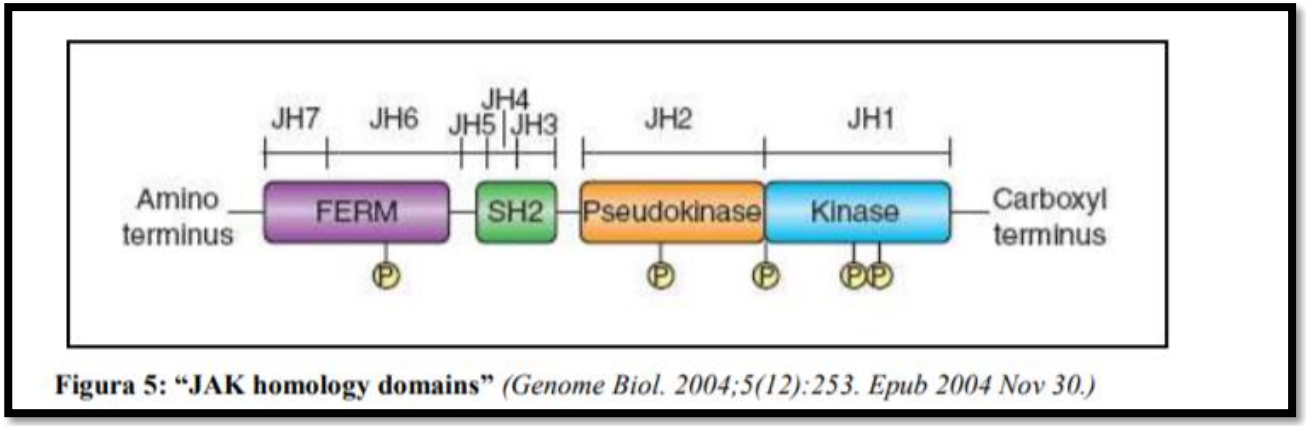
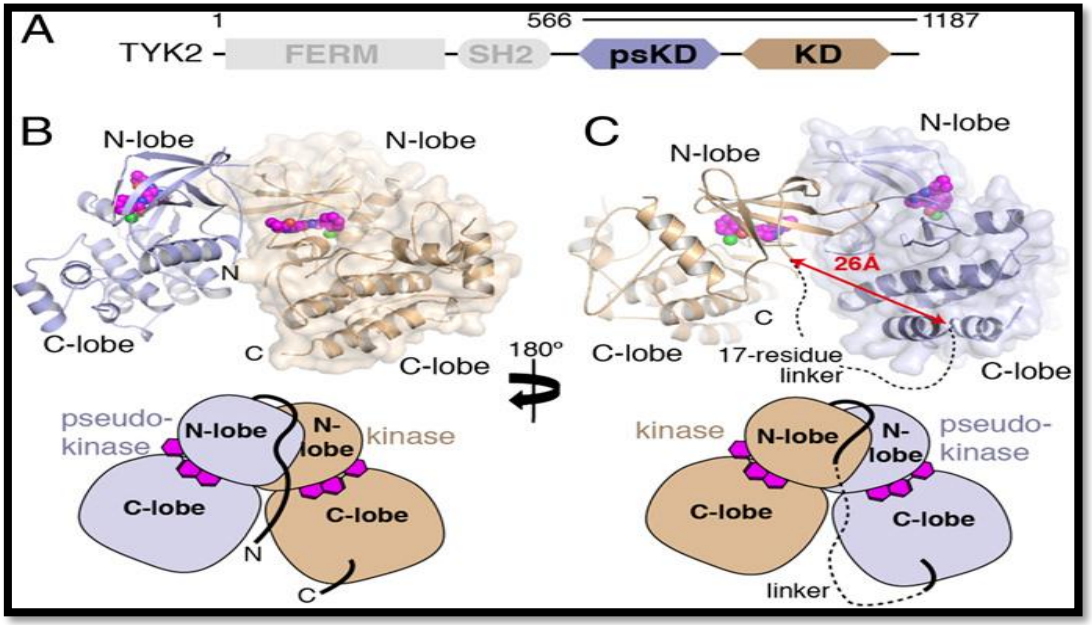
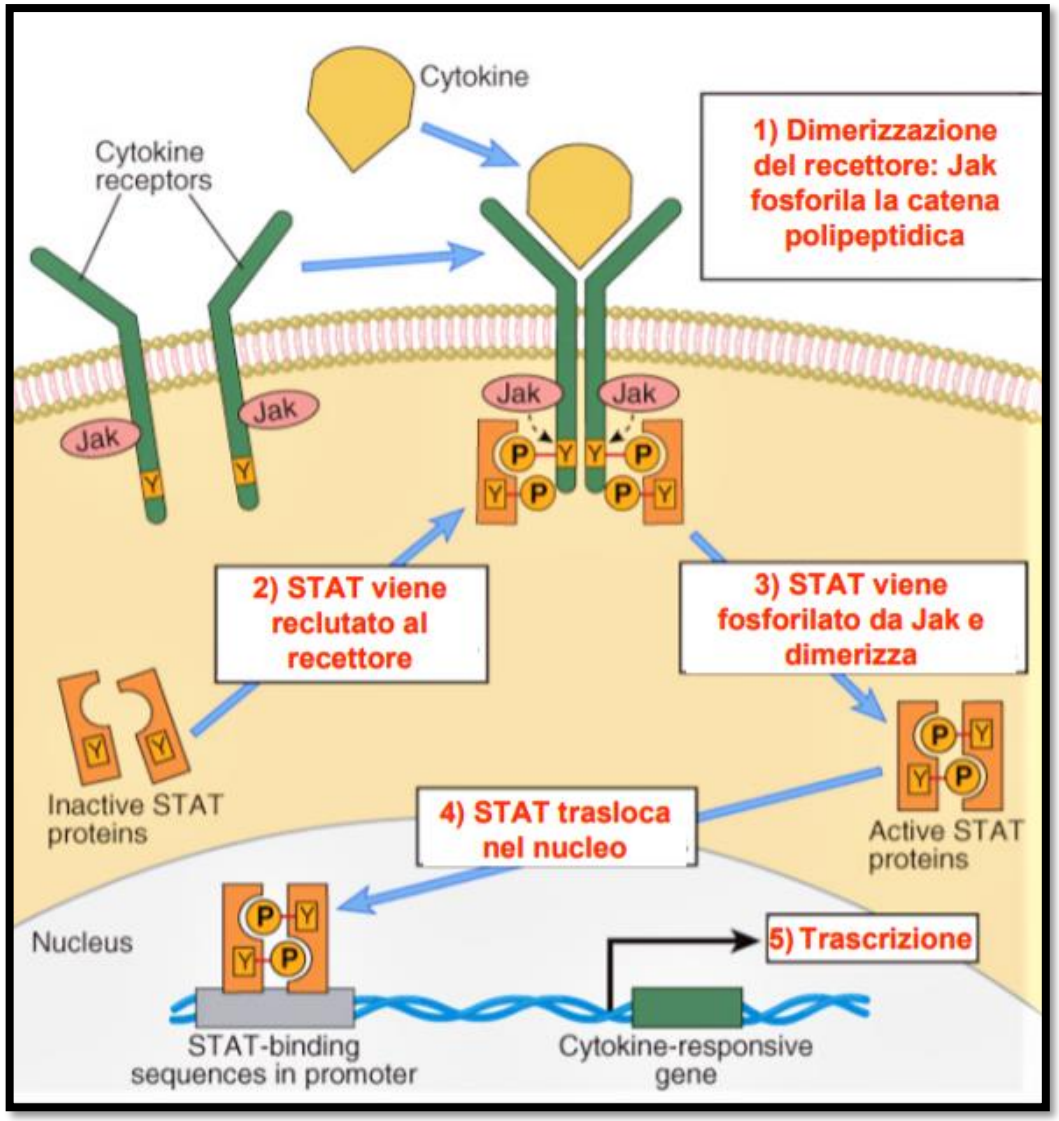


Figura 5: "JAK homology domains" *(Genome Biol. 2004;5(12):253. Epub 2004 Nov 30.)*

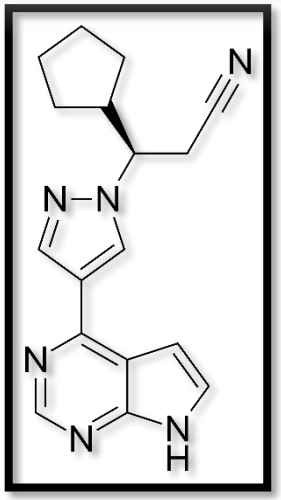


Pathway jak-stat

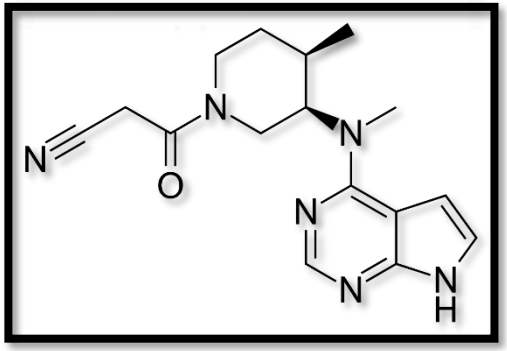


Inibitori delle JAK di prima generazione

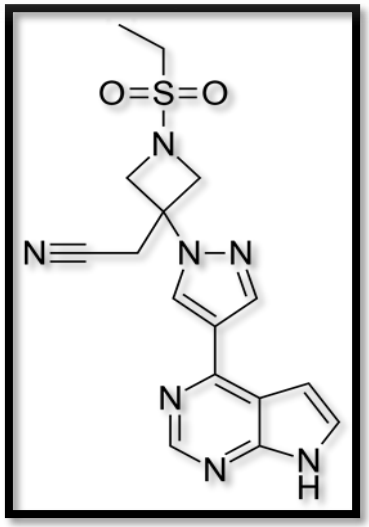
Ruxolitinib
FDA approved 2011



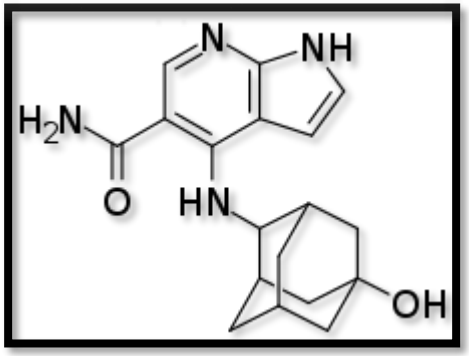
Tofacitinib
FDA approved 2012



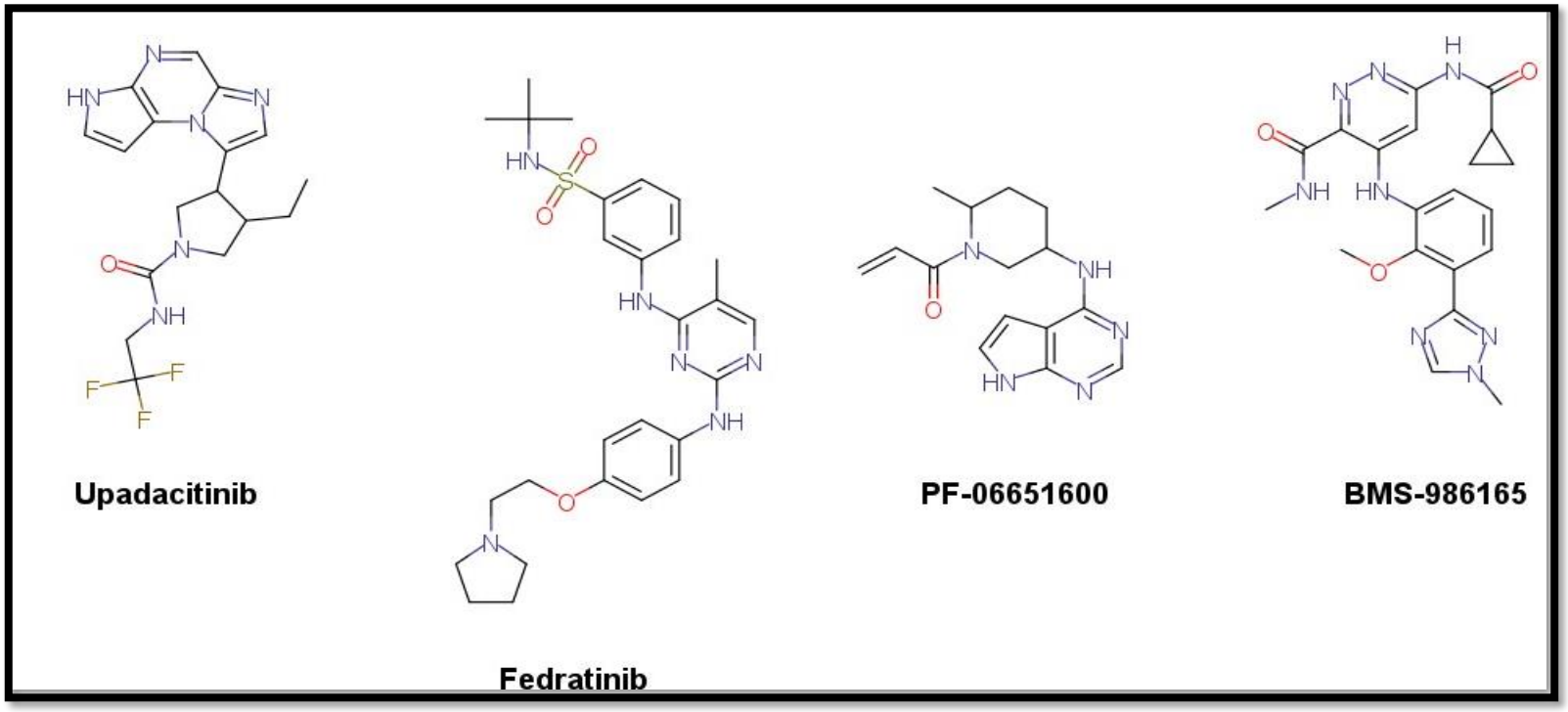
Baricitinib
FDA approved 2018



Peficitinib
FDA approved 2019



Inibitori JAK non selettivi



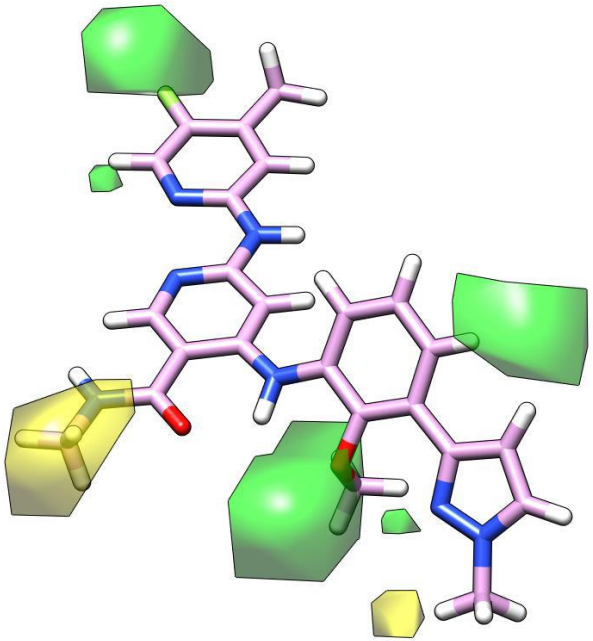
Inibitori JAK selettivi

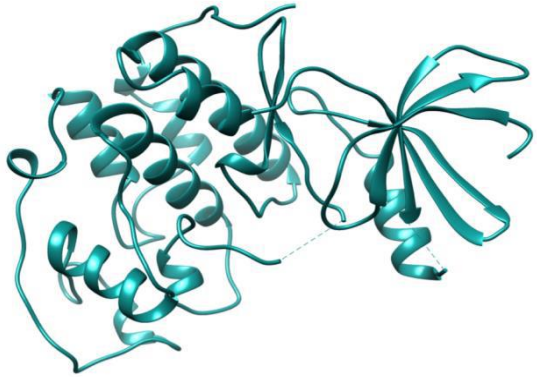
Approccio ligand-based e structure-based

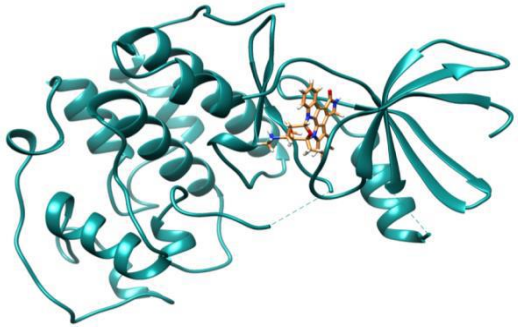
DRUG DESIGN

LIGAND-BASED

STRUCTURE-BASED



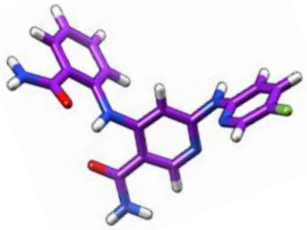




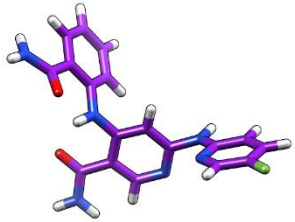
Dataset

SMILES	IC50(nM)
NC(=O)C1=CN=C2N1N=C(NC1=CC(=CC(=C1)C#N)C(F)(F)F)C=C2NC1CC1	480
NC(=O)C1=CN=C2N1N=C(NC1=CC=CC=C1)C=C2NC1CC1	32
CC1=CC(NC2=NN3C(=CN=C3C(NC3CC3)=C2)C(N)=O)=CC(C)=C1	22
CC1=CC(F)=CC(NC2=NN3C(=CN=C3C(NC3CC3)=C2)C(N)=O)=C1	24
NC(=O)C1=CN=C2N1N=C(NC1CCCCC1)C=C2NC1CC1	220
CNC1=CC(NC2=CC(C)=CC(C)=C2)=NN2C(=CN=C12)C(N)=O	7
CCNC1=CC(NC2=CC(C)=CC(C)=C2)=NN2C(=CN=C12)C(N)=O	17
CC1=CC(NC2=NN3C(=CN=C3C(NC3CCC3)=C2)C(N)=O)=CC(C)=C1	110

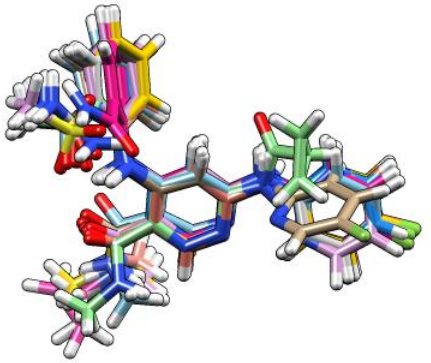
Conversione in 3D delle molecole



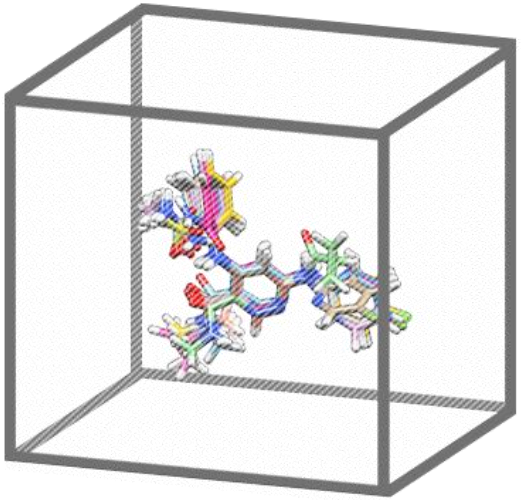
Analisi Conformazional



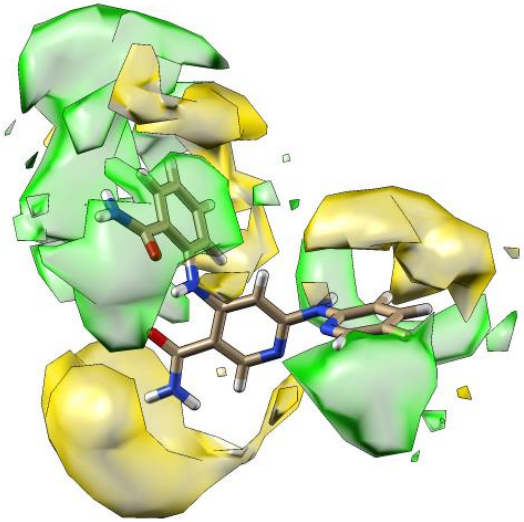
Allineamento



Calcolo dei MIF e generazione del modello matematico-statistico



Analisi grafica



LIGAND-BASED

1. Identification of imidazo[1,2-b]pyridazine TYK2 pseudokinase ligands as potent and selective allosteric inhibitors of TYK2 signalling (Moslin et al., 2017)

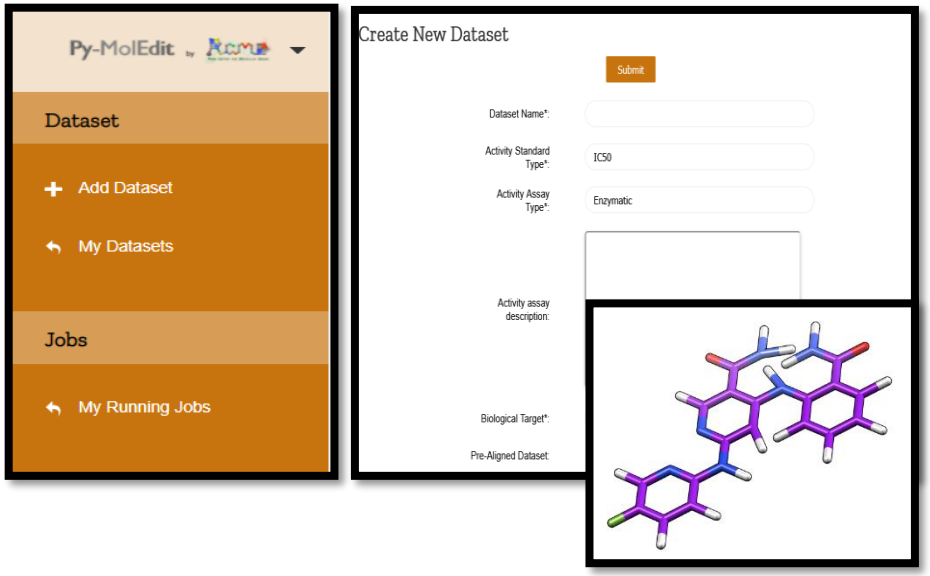
2. Identification of N-Methyl Nicotinamide and N-Methyl Pyridazine-3-Carboxamide Pseudokinase Domain Ligands as Highly Selective Allosteric Inhibitors of Tyrosine Kinase 2 (TYK2) (Moslin et al., 2019)

3. Highly Selective Inhibition of Tyrosine Kinase 2 (TYK2) for the Treatment of Autoimmune Diseases: Discovery of the Allosteric Inhibitor BMS-986165 (Wrobleski et al., 2019)

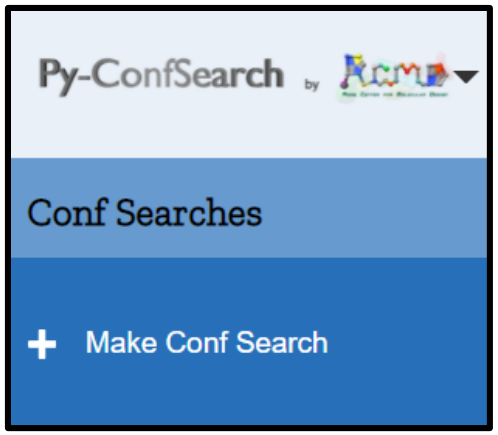
4. Discovery of BMS-986202: A Clinical Tyk2 Inhibitor that Binds to Tyk2 JH2 (Liu et al., 2021)

Dataset

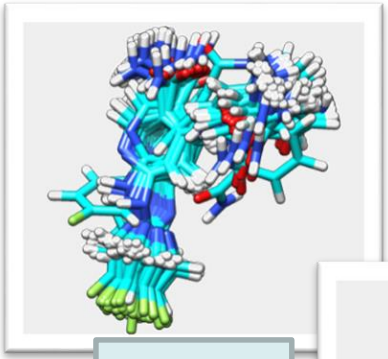
SMILES	IC50(nM)
NC(=O)C1=CN=C2N1N=C(NC1=CC(=CC(=C1)C#N)C(F)(F)F)C=C2NC1CC1	480
NC(=O)C1=CN=C2N1N=C(NC1=CC=CC=C1)C=C2NC1CC1	32
CC1=CC(NC2=NN3C(=CN=C3C(NC3CC3)=C2)C(N)=O)=CC(C)=C1	22
CC1=CC(F)=CC(NC2=NN3C(=CN=C3C(NC3CC3)=C2)C(N)=O)=C1	24
NC(=O)C1=CN=C2N1N=C(NC1CCCCC1)C=C2NC1CC1	220
CNC1=CC(NC2=CC(C)=CC(C)=C2)=NN2C(=CN=C12)C(N)=O	7
CCNC1=CC(NC2=CC(C)=CC(C)=C2)=NN2C(=CN=C12)C(N)=O	17
CC1=CC(NC2=NN3C(=CN=C3C(NC3CCC3)=C2)C(N)=O)=CC(C)=C1	110



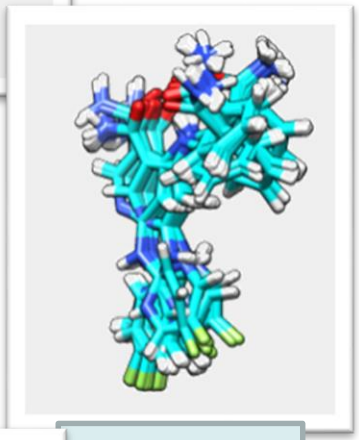
Analisi conformazionale



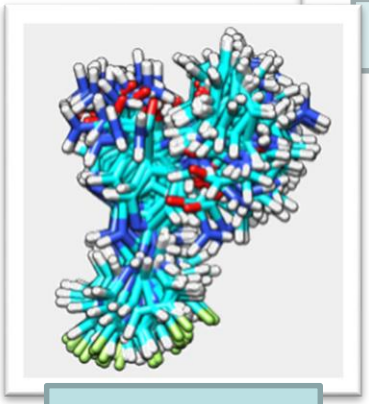
- Balloon → EEM, MMFF94, SFKEEM
- Open Babel → UFF, MMFF94S, MMFF94, GAFF, GHEMICAL
- RDKIT → UFF, MMFF94S, MMFF94



Balloon



RDkit



Openbabel

Conf Search ID	Status		Method	Max Confs	CG MaxIters	SD MaxIters	Force Field
4748	completed	Results	openbabel	50	50	50	UFF
4749	completed	Results	openbabel	50	50	50	MMFF94
4750	completed	Results	openbabel	50	50	50	MMFF94s
4887	completed	Results	openbabel	50	50	50	GAFF
4888	completed	Results	openbabel	50	50	50	Ghemical
4889	completed	Results	balloon	50	50	50	EEM
4890	completed	Results	balloon	50	50	50	MMFF94
4891	completed	Results	balloon	50	50	50	SFKEEM
4892	completed	Results	rdkit	50	50		UFF
4893	completed	Results	rdkit	50	50		MMFF94
4894	completed	Results	rdkit	50	50		MMFF94s

Allineamenti

Py-Align by Acmd

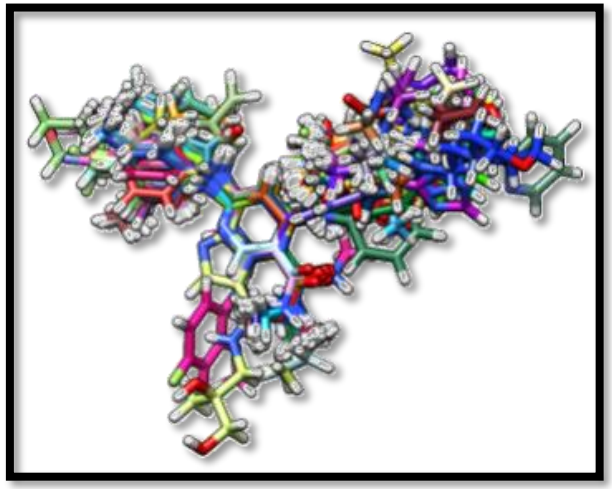
RDKIT

SHAEP

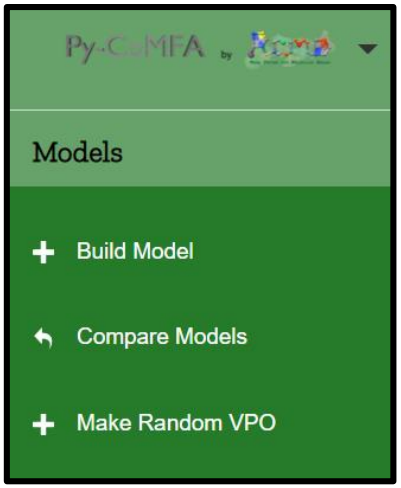
SIMILARITY ONLY SHAPE

BEST SCORE
LOWEST RMSD
TANIMOTO DISTANCE
PROTUDE DISTANCE

1. Meno attiva
2. Più attiva
3. Più flessibile
4. Più rigida
5. Più pesante
6. Più lunga
7. Meno polare
8. Più polare
9. Con più alta MR
10. Con più bassa MR
11. Con il più alto numero HA
12. Con il più basso numero di HA
13. Con il più alto numero HD
14. Con il più basso numero di HD
15. Con il più alto LogP
16. Con il più basso LogP



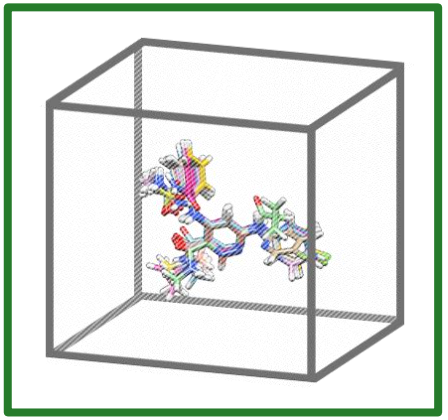
Modelli ligand-based



Parametri di default:

- Probe = C.3
- Grid Spacing = 2 Å
- Grid Extension = 5 Å
- Dielectric costant = 8
- Min/Max energy of cutoff value = 30 kcal/mol
- Minimum sigma = 0.05

MIF (MOLECULAR INTERACTION FIELD)



$$q^2 = 1 - \frac{\sum_{i=1}^{n} (y_{exp\_i} - y_{pred\_i})^2}{\sum_{i=1}^{n} (y_{exp\_i} - \bar{y})^2}$$

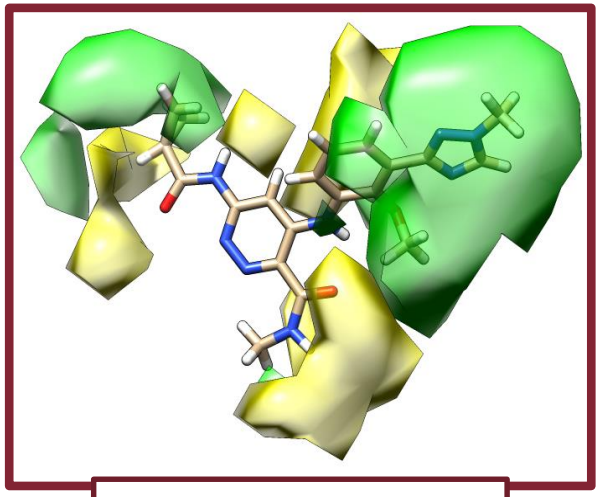
Ottimizzazione tramite VPO:

- Probe = C.3.H3, H.P, N.am, O.2
- Grid Spacing = 1, 3.5, 0.1
- Grid Extension = 1, 10, 1
- Dielectric costant = 1, 80, 1
- Min/Max energy of cutoff value = 5, 35, 1
- Minimum sigma = 0.1, 1.0, 0.1

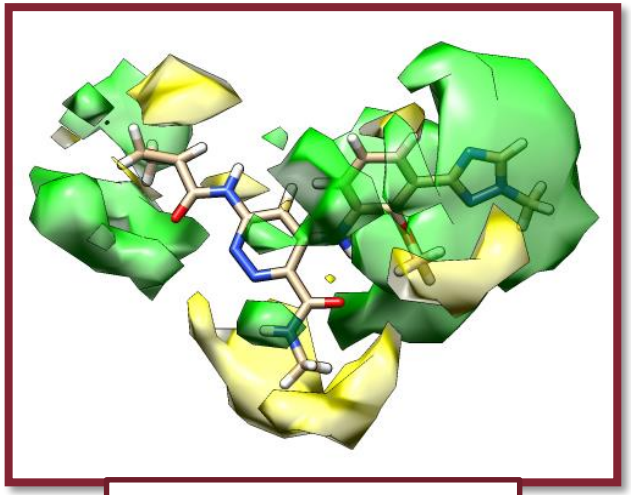
Risultati

RDkit	r^2	q^2	OPC	P	GS	GE	ε	CO	MS
Tyk2_inhibitor_M	0.883	0.563	5	C.3.H3	1.3	7	13	5	0.5
Tyk2_inhibitor_S	0.915	0.640	6	N.am	1.5	2	37	15	1.0
Tyk2_inhibitor_C	0.815	0.570	4	C.3.H3	2.4	1	69	16	0.4

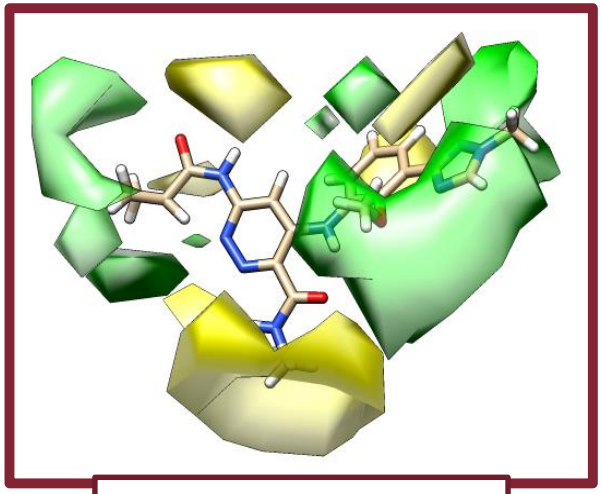
$r^2 = 0.871 \pm 0.051$
$q^2 = 0.591 \pm 0.043$



TYK2_inhibitor_M



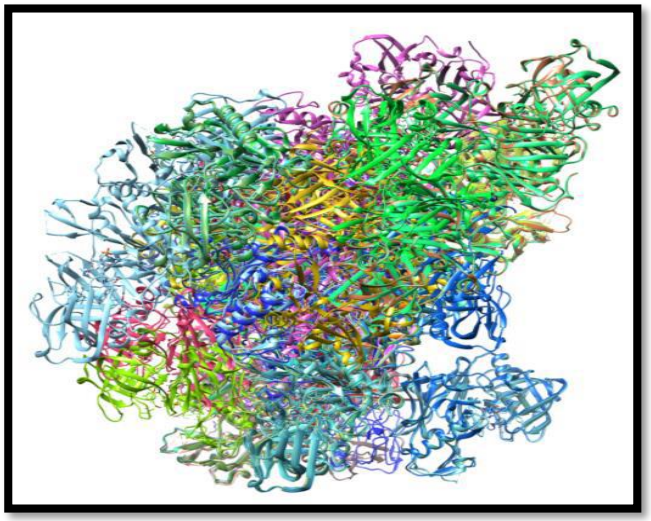
TYK2_inhibitor_S

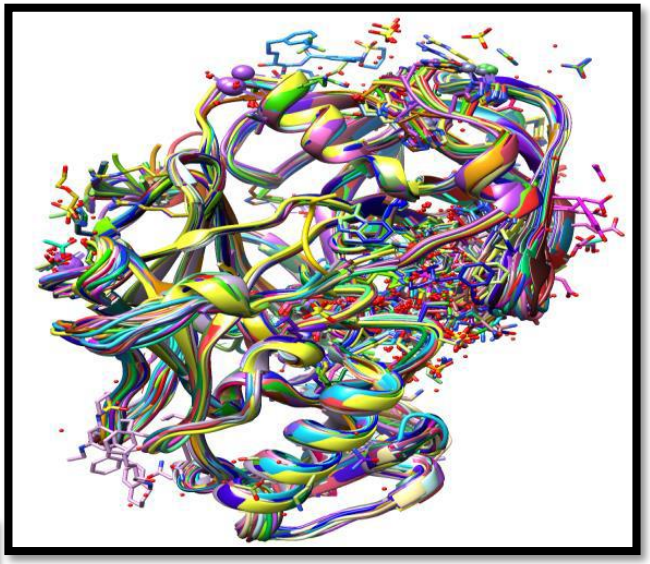


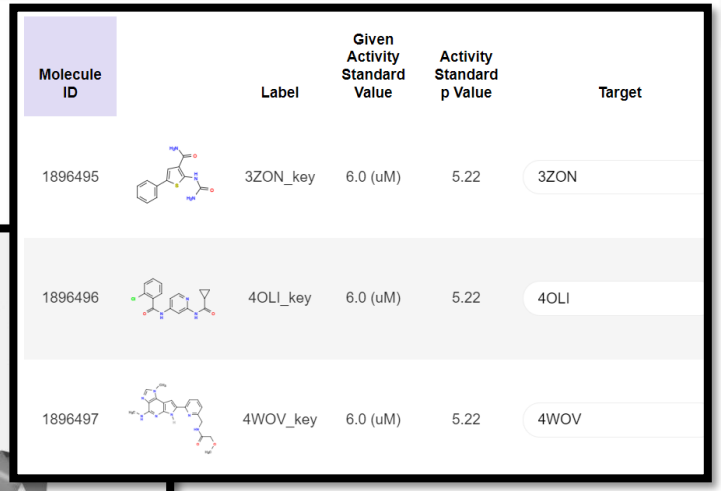
TYK2_inhibitor_C

Approccio structure-based

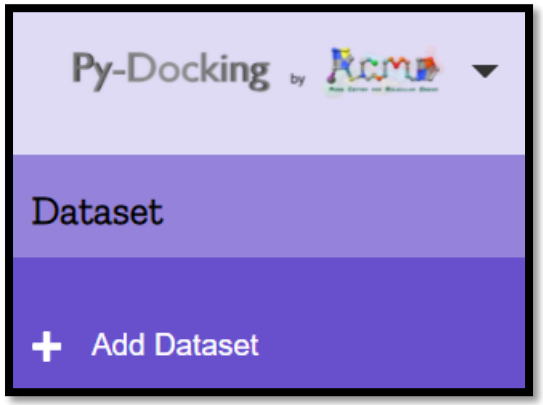
1. 3ZON
2. 4OLI
3. 4WOV
4. 5C01
5. 5C03
6. 5TKD
7. 6NSL
8. 6NZE
9. 6NZF
10. 6NZH
11. 6NZP
12. 6NZQ
13. 6NZR
14. 7AX4

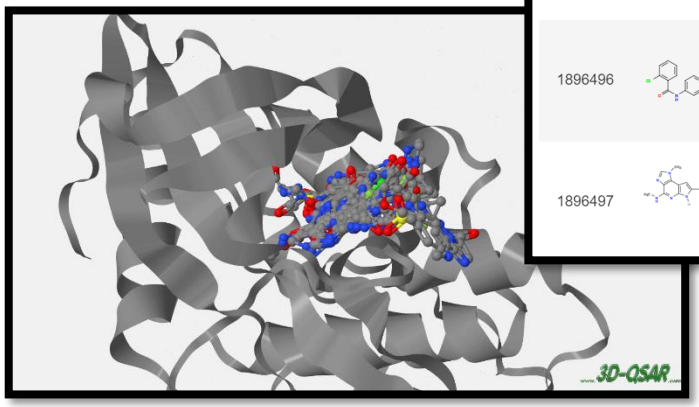






Molecule ID		Label	Given Activity Standard Value	Activity Standard p Value	Target
1896495		3ZON_key	6.0 (uM)	5.22	3ZON
1896496		4OLI_key	6.0 (uM)	5.22	4OLI
1896497		4WOV_key	6.0 (uM)	5.22	4WOV





Docking

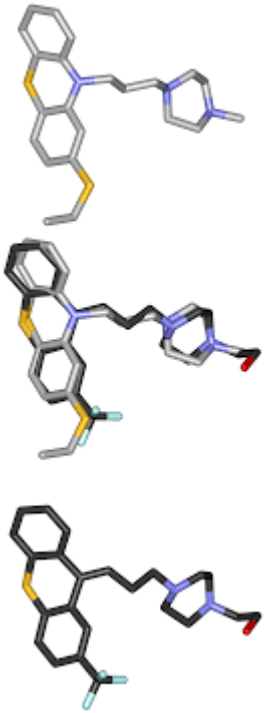
DOCKING

PLANTS:

Plp
Plp95
Chemplp

SMINA:

Vina/minim
Vinardo/minim
Ad4scoring/minim



ECRD (Experimental Conformation Re-Docking)
RCRD (Random Conformation Re-Docking)

ECCD (Experimental Conformation Cross-Docking)
RCCD (Random Conformation Re-Docking)

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

$$DA = frmsd \leq a + 0.5\,(frmsd \leq b - frmsd \leq a)$$

DA %

METHOD/SCORING FUNCTION		ECRD	RCRD	ECCD	RCCD
Plants Chemplp	MEDIA	75,42	78,57	44,04	42,86
	DS	±11,47	±6,18	±4,13	±7,15
Plants Plp	MEDIA	58,85	69,05	44,04	46,43
	DS	±7,73	±10,31	±7,44	±6,18
Plants Plp95	MEDIA	71,42	72,62	46,76	46,42
	DS	±0,01	±2,06	±5,92	±0,01
Smina ad4scoring	MEDIA	51,19	54,76	46,42	45,23
	DS	±5,46	±8,99	±0,01	±4,13
Smina ad4scoring min.	MEDIA	8,33	16,66	1,19	2,38
	DS	±5,46	±5,45	±2,06	±2,06
Smina Vina	MEDIA	35,71	39,28	35,71	34,52
	DS	±0,00	±3,57	±0,00	±4,12
Smina Vina min.	MEDIA	9,52	8,33	0,00	4,76
	DS	±2,06	±2,06	±0,00	±4,12
Smina Vinardo	MEDIA	31,76	34,52	29,76	36,90
	DS	±14,92	±7,43	±8,99	±8,99
Smina Vinardo min.	MEDIA	2,38	9,52	7,14	2,38
	DS	±4,12	±4,12	±7,14	±4,12

Conclusioni

Nonostante la casualità insita negli algoritmi di calcolo :

1. **Nonostante l'alta intrinseca casualità nei calcoli (conversione SMILES --> 3D, Analisi conformazionale, VPO) i modelli 3D-QSAR conducono alle stesse conclusioni**
2. **La deviazione standard dei valori di q2 dei modelli ottenuti partendo dalle stesse analisi conformazionali risulta essere molto bassa**
3. **Il metodo di docking più accurato è risultato concordante per le tre applicazioni in parallelo**
4. **La DA ottenuta ha mostrato una minima variazione.**
5. **Questa procedura di calcolo è efficace e riproducibile e in linea di massima potrebbe essere applicata a qualsiasi target**
6. **Lo studio sarà ampliato:**
   - **Applicazione ad un test set esterno per valutare la predittività**
   - **Derivazione di modelli SB (3D-QSAR e COMBINE)**

Dai risultati possiamo affermare che la procedura applicata mediante il portale 3d-qsar.com permette di definire modelli altamente affidabili e riproducibili.