SBDD

SAPIENZA
UNIVERSITÀ DI ROMA

Download Ligand Scout

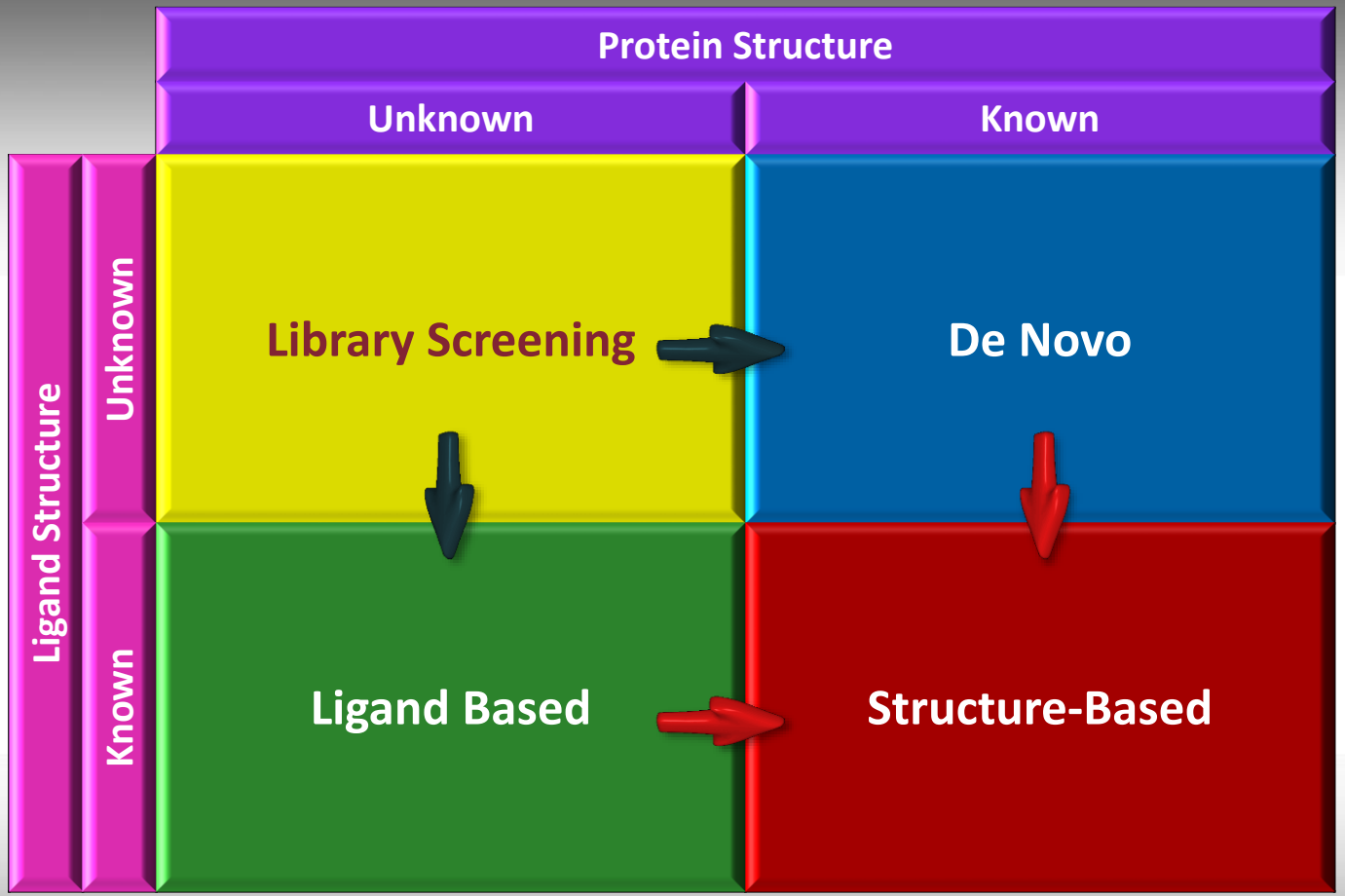
Linux:
https://www.inteligand.com/ligandscout4/downloads/LigandScout_4_4_5_linux64_20200714.tar.gz

Windows:
https://www.inteligand.com/ligandscout4/downloads/LigandScout_4_4_5_win64_20200714.exe

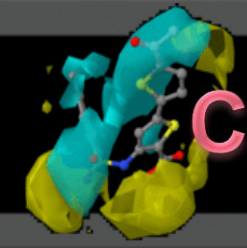
Mac:
https://www.inteligand.com/ligandscout4/downloads/LigandScout_4_4_5_macos_20200714.dmg

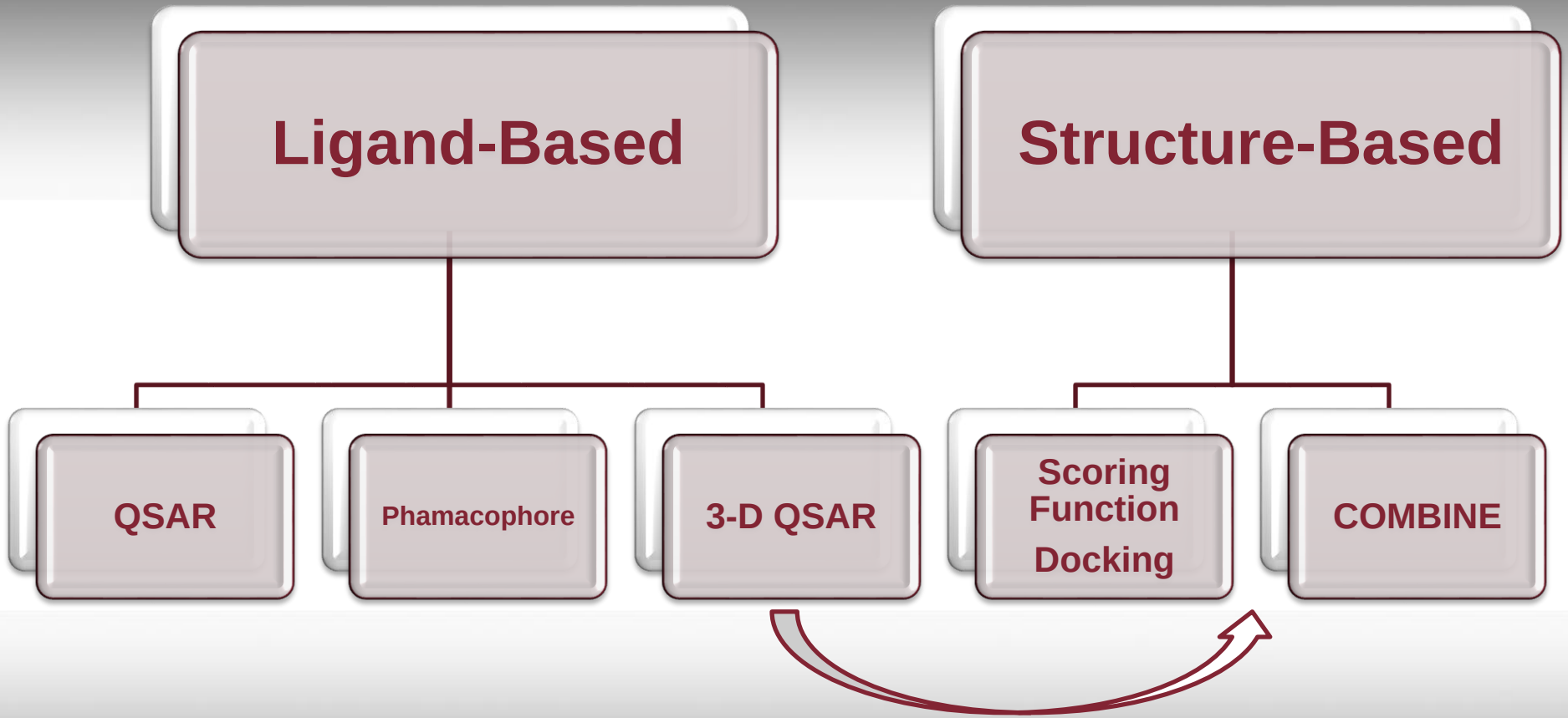
LICENZA:

6713013763892254839 8


Protein Structure
Unknown
Known
Ligand Structure
Unknown
Known
Library Screening
De Novo
Ligand Based
Structure-Based

CADD methods in Drug Design





SBDD methods

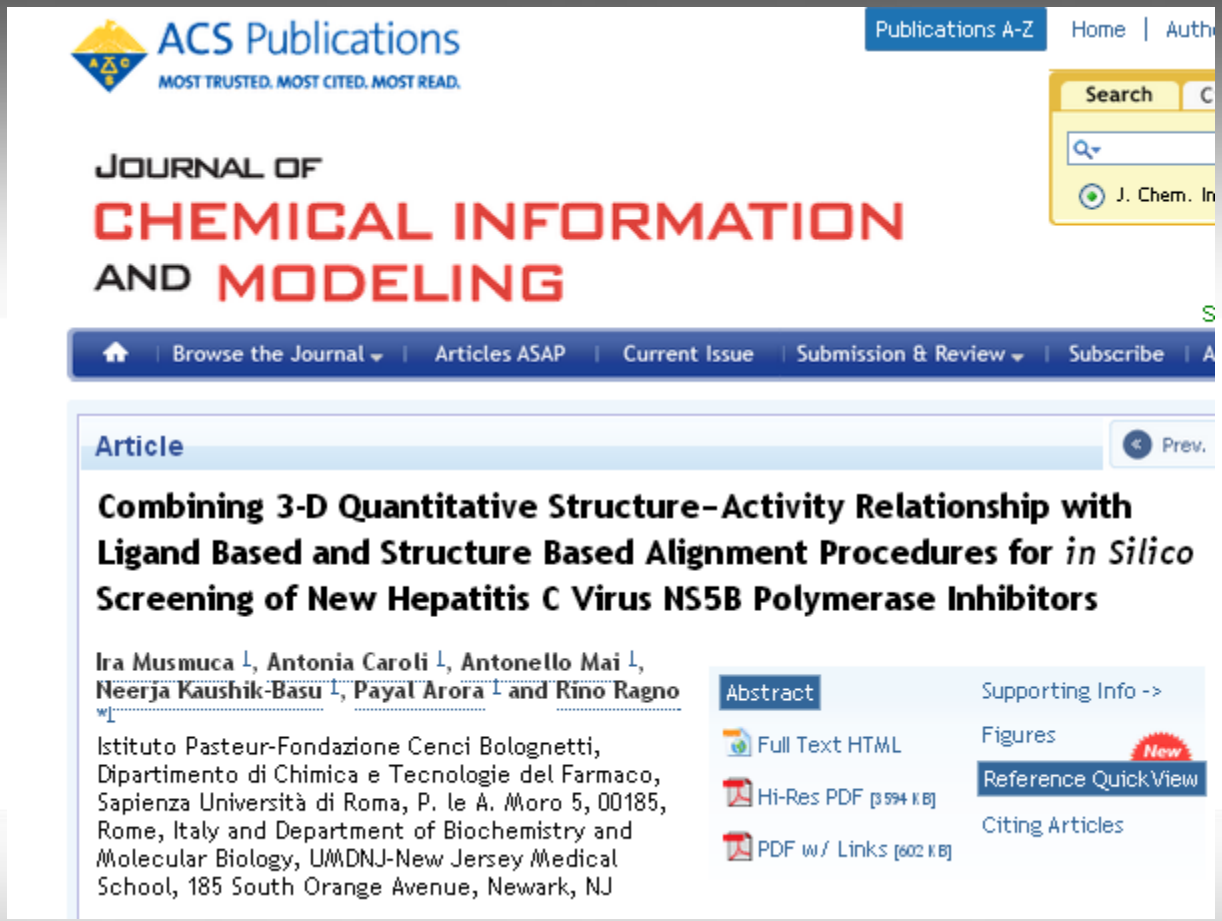
Structure-Based Drug Design Approaches

- 3-D QSAR (Molecules aligned by SB methods)
- Scoring Function (A Sort of QSARs for Molecular Docking)
- COMBINE (A Full SB 3-D QSAR)
- Proteochemometric (A QSAR SB derived method)
- Molecular Dynamics
- Pharmacophoric Approaches (Using Targets Structures)
- Homology Modeling (To Build Target Structures)

SB 3-D QSAR

3-D QSAR models developed with molecular alignment rules employing 3-D structure of the target

SB 3-D QSAR



ACS Publications
MOST TRUSTED. MOST CITED. MOST READ.

Publications A-Z | Home | Auth

Search | C

J. Chem. In

JOURNAL OF
CHEMICAL INFORMATION
AND MODELING

Browse the Journal | Articles ASAP | Current Issue | Submission & Review | Subscribe | A

Article

Prev.

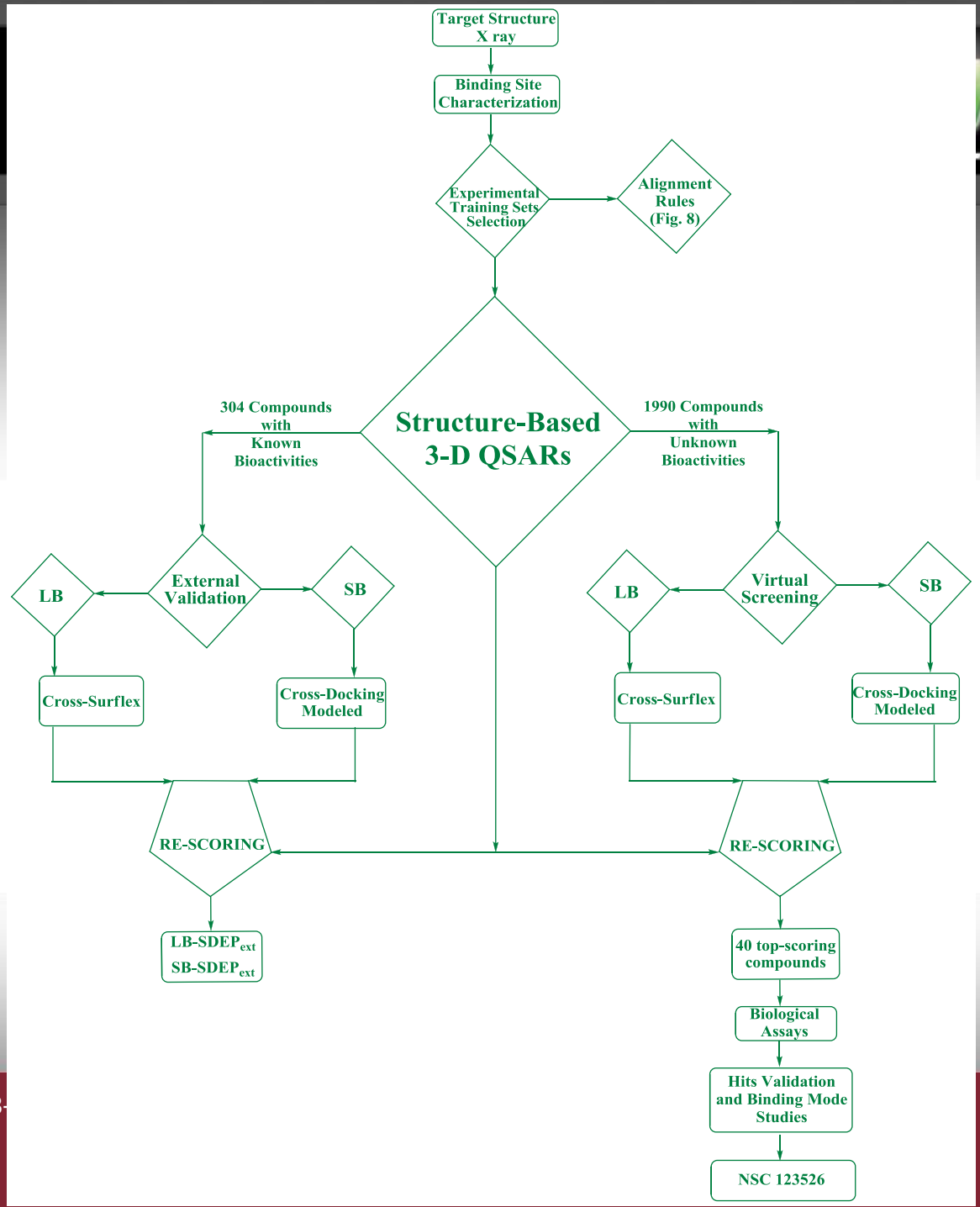
Combining 3-D Quantitative Structure–Activity Relationship with Ligand Based and Structure Based Alignment Procedures for *in Silico* Screening of New Hepatitis C Virus NS5B Polymerase Inhibitors

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

Abstract | Full Text HTML | Hi-Res PDF [3594 KB] | PDF w/ Links [602 KB]

Supporting Info -> | Figures | Reference QuickView (New) | Citing Articles


Target Structure X ray
Binding Site Characterization
Experimental Training Sets Selection
Alignment Rules (Fig. 8)
Structure-Based 3-D QSARs
304 Compounds with Known Bioactivities
1990 Compounds with Unknown Bioactivities
LB
External Validation
SB
LB
Virtual Screening
SB
Cross-Surflex
Cross-Docking Modeled
Cross-Surflex
Cross-Docking Modeled
RE-SCORING
RE-SCORING
LB-SDEP$_{ext}$
SB-SDEP$_{ext}$
40 top-scoring compounds
Biological Assays
Hits Validation and Binding Mode Studies
NSC 123526

CoMFA/3-D QSAR Procedure

Training Set

Molecular Alignment

Molecular Interaction Fields (MIF)

Model Generation

Internal and External Validation

CoMFA/3-D QSAR Procedure

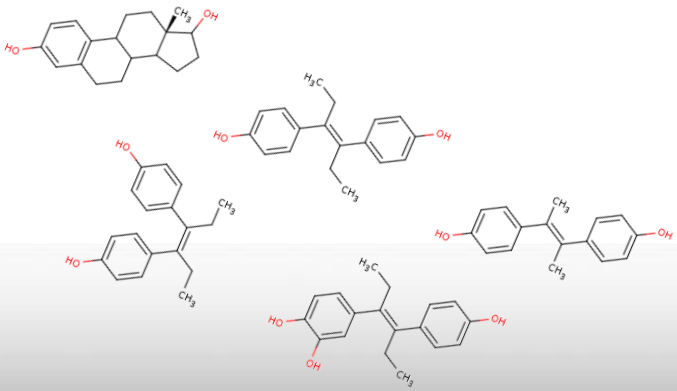
Training Set

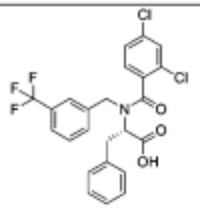
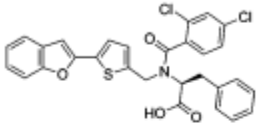
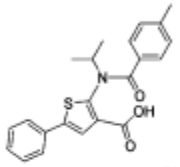
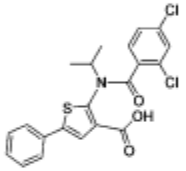
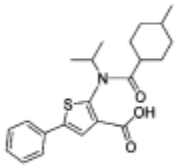
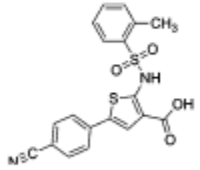
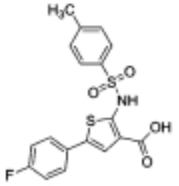
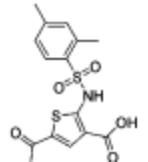
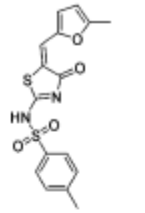
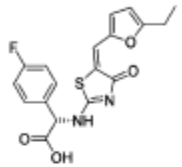
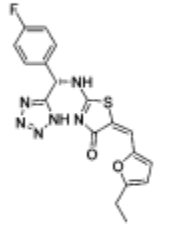
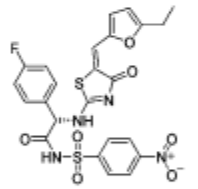
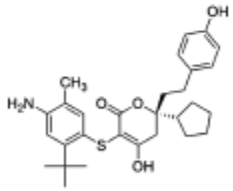
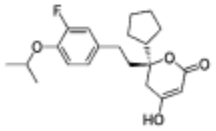
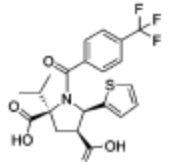
Molecular lignment

Molecular Interaction Fields (MIF)

Model Generation

Internal and External Validation



PDB Code	Ligand Name	Entry	Ligand Structure	IC_{50} (μM)
1NHU	(2*S*)-2-[(2,4-dichloro-benzoyl)-(3-trifluoromethyl-benzyl)-amino]-3-phenyl-propionic acid	**1**		1.7
1NHV	(2*S*)-2-[(5-benzofuran-2-yl-thiophen-2-yl methyl)-(2,4-dichloro-benzoyl)-amino]-3-phenyl-propionic acid	**2**		8.6
1YVX	3-[isopropyl(4-methylbenzoyl)amino]-5-phenylthiophene-2-carboxylic acid	**3**		8.0
1YVZ	3-[(2,4-dichlorobenzoyl) (isopropyl)amino]-5-phenylthiophene-2-carboxylic acid	**4**		4.4
2GIR	3-{isopropyl[(trans-4-methylcyclohexyl)carbonyl]amino}-5-phenylthiophene-2-carboxylic acid	**5**		1.5
2D3U	5-(4-cyanophenyl)-3-{[(2-methylphenyl)sulfonyl] amino}thiophene-2-carboxylic acid	**6**		0.2663
2D3Z	5-(4-fluorophenyl)-3-{[(4-methylphenyl)sulfonyl] amino}thiophene-2-carboxylic acid	**7**		0.2918
2D41	5'-acetyl-4-{[(2,4-dimethylphenyl)sulfonyl] amino}-2,2'-bithiophene-5-carboxylic acid	**8**		0.3073
2HWH	4-methyl-N-{(5*E*)-5-[(5-methyl-2-furyl)methylene]-4-oxo-4,5-dihydro-1,3-thiazol-2-yl} benzenesulfonamide	**9**		2.0
2HWI	(2*S*)-({(5*Z*)-5-[(5-ethyl-2-furyl)methylene]-4-oxo-4,5-dihydro-1,3-thiazol-2-yl}amino)(4-fluorophenyl)acetic acid	**10**		3.0
2I1R	(5*Z*)-5-[(5-ethyl-2-furyl)methylene]-2-{[(*S*)-(4-fluorophenyl)(1H-tetrazol-5-yl)methyl] amino}-1,3-thiazol-4(5H)-one	**11**		9.7
2O5D	(2*S*)-2-({(5*Z*)-5-[(5-ethyl-2-furyl)methylene]- 4-oxo-4,5-dihydro-1,3-thiazol-2-yl}amino)- 2-(4-fluorophenyl)-N-[(4-nitrophenyl)sulfonyl] acetamide	**12**		7.0
1OS5	3-(4-amino-2-tert-butyl-5-methyl-phenylsulfanyl)- 6-cyclopentyl-4-hydroxy-6-[2-(4-hydroxy- phenyl)- ethyl]-5,6-dihydro-pyran-2-one	**13**		0.93
2HAI	(6*S*)-6-cyclopentyl-6-[2-(3-fluoro-4-isopropoxyphenyl)ethyl]- 4-hydroxy-5,6-dihydro-2H-pyran-2-one	**14**		0.53
2JC0	(2*S*,4*S*,5*R*)-2-isobutyl-5-(2-thienyl)-1-[4-(trifluoromethyl)benzoyl]pyrrolidine-2,4-dicarboxylic acid	**15**		20.0

CoMFA/3-D QSAR Procedure

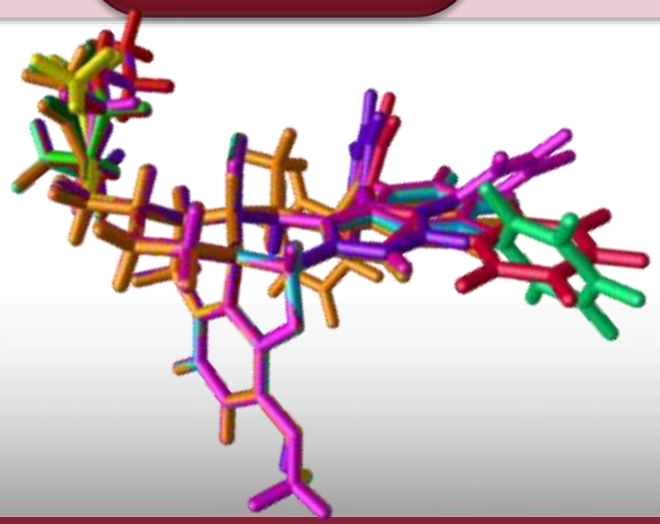
Training Set

Molecular Alignment

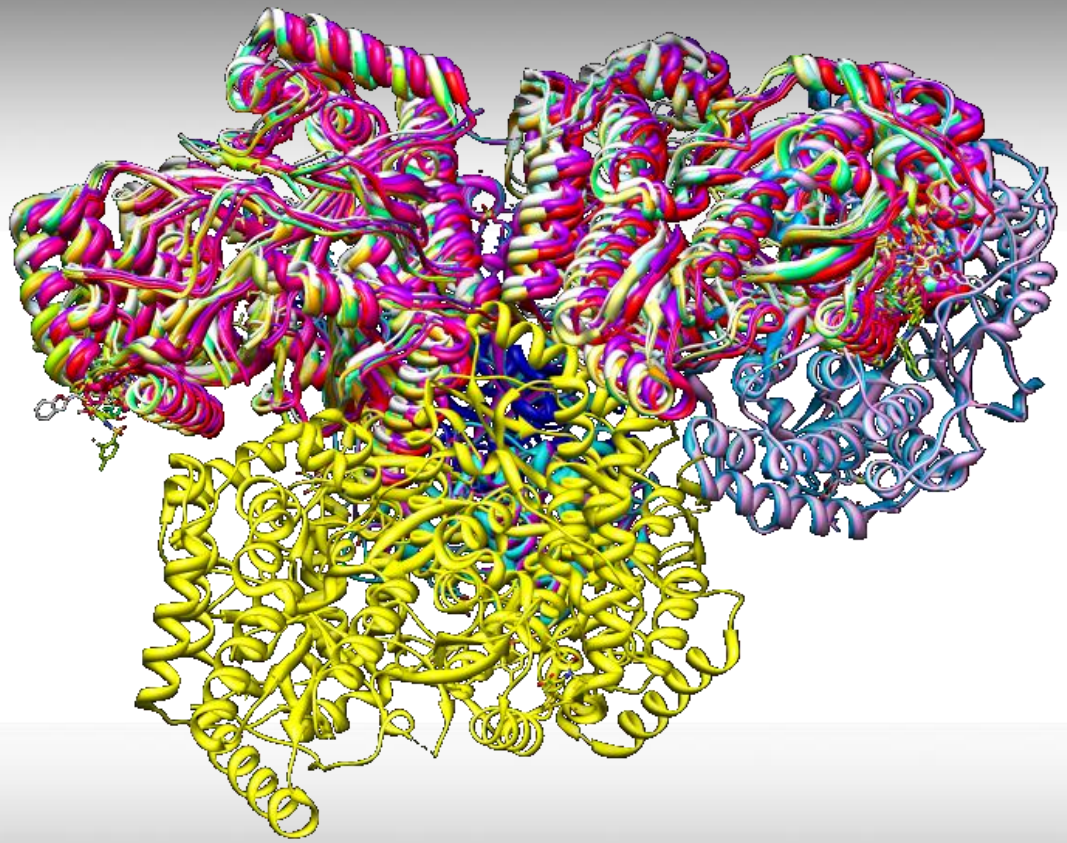
Molecular Interaction Fields (MIF)

Model Generation

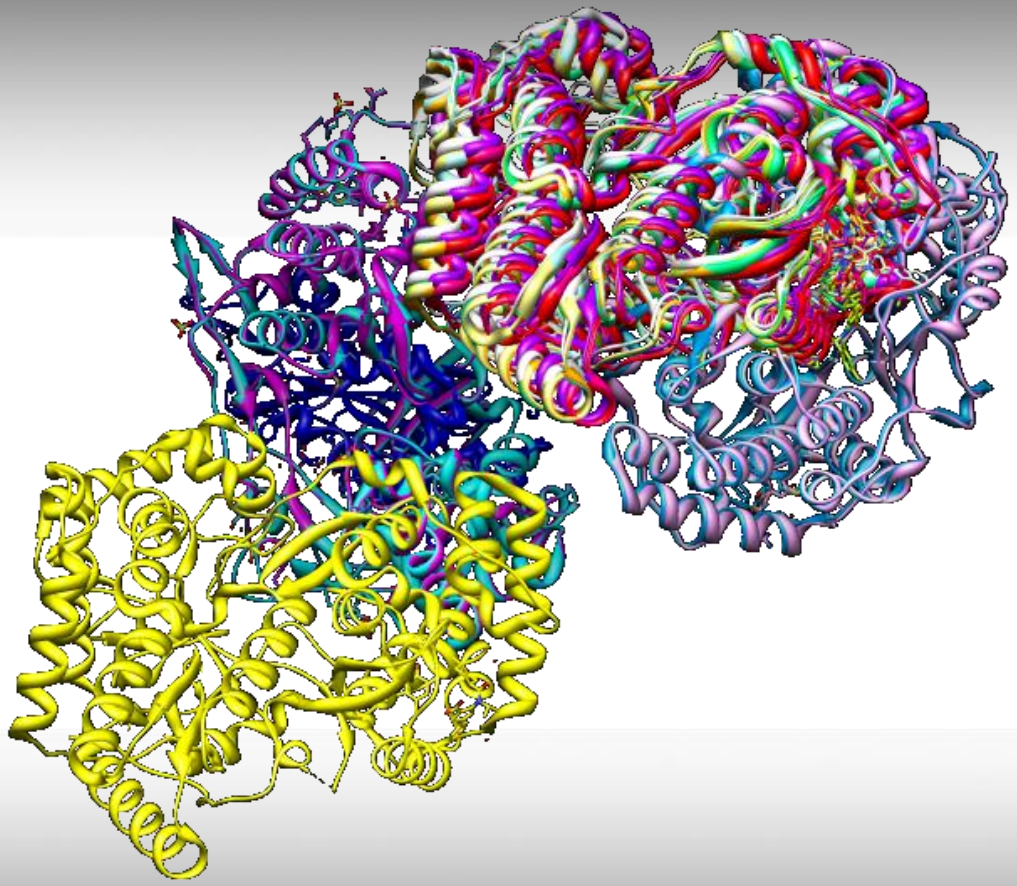
Internal and External Validation



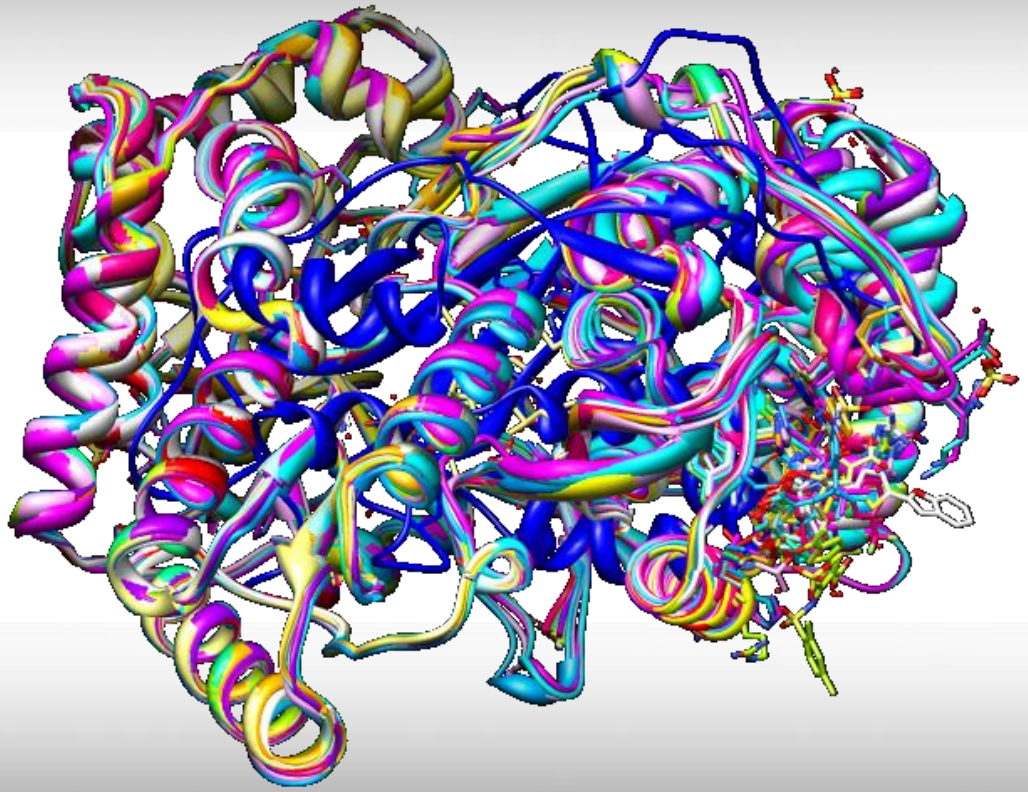
SB Alignment



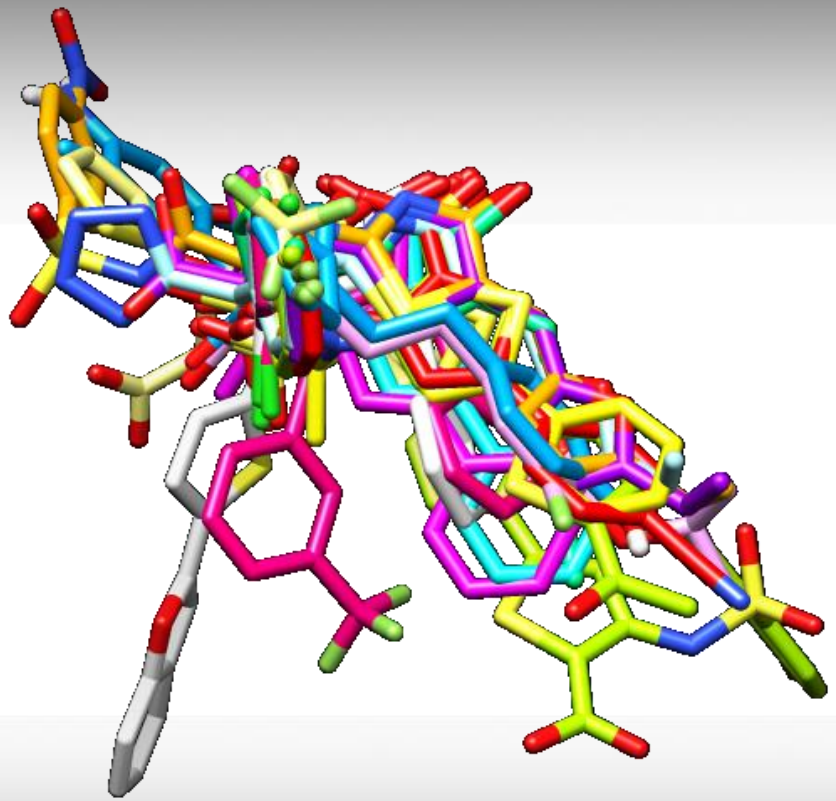
SB Alignment



SB Alignment



SB Alignment



CoMFA/3-D QSAR Procedure

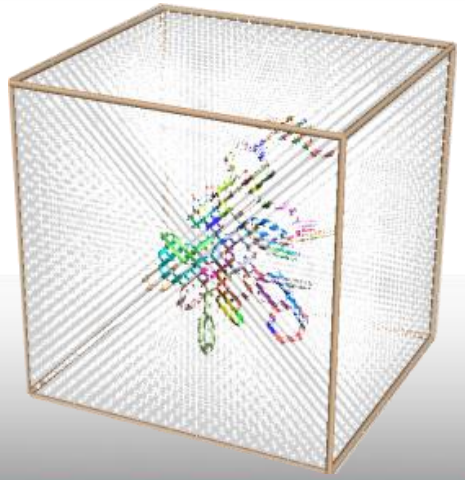
Training Set

Molecular Alignment

Molecular Interaction Fields (MIF)

Model Generation

Internal and External Validation



CoMFA/3-D QSAR Procedure

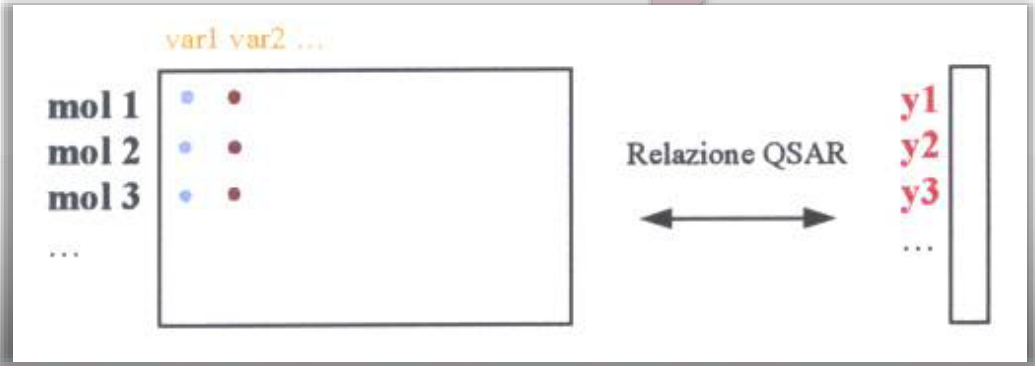
Training Set

Molecular lignment

Molecular Interaction Fields (MIF)

Model Generation

Internal and External Validation



CoMFA/3-D QSAR Procedure

Training Set

Molecular Alignment

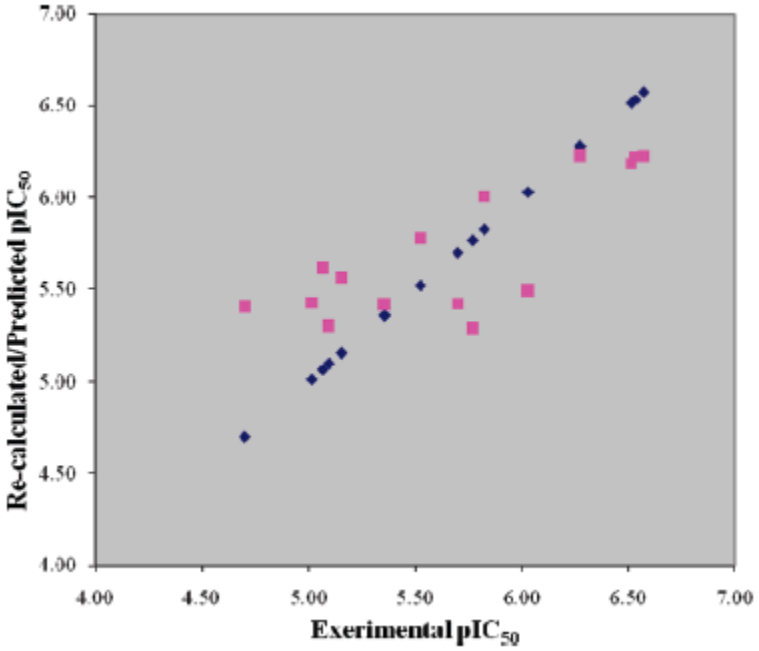
Molecular Interaction Fields (MIF)

Model Generation

Internal and External Validation

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

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



model	N	GRID probe	V	PC	r^2	q^2	$SDEP_{CV\text{-}LSO}$
thumb training set	15	C1=	5133	3	0.99	0.69	0.31
palm training set	10	C1=	3848	3	0.99	0.55	0.66

[a] *N*, number of compounds in the training set; *V*, number of GOLPE variables; PC, optimal number of principal components; r^2, conventional square correlation coefficient; q^2, cross-validation correlation coefficient; SDEP, cross-validated standard error of prediction using the leave-five-out cross-validation method.

CoMFA/3-D QSAR Procedure

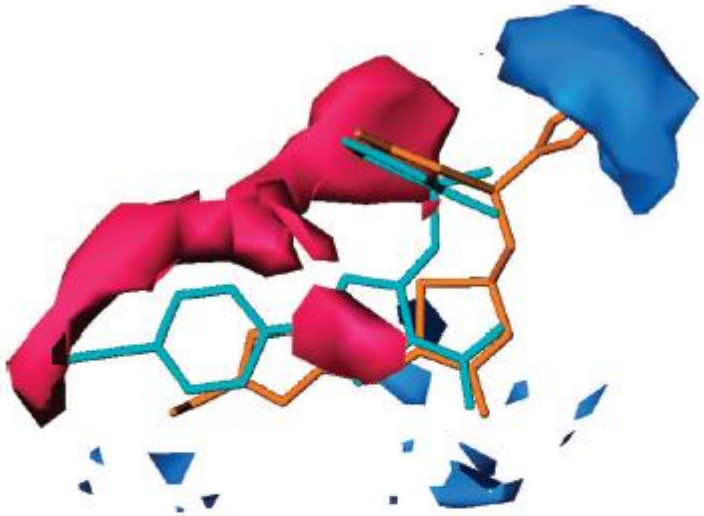


Figure 3. Contour maps of the PLS coefficients derived from C1= GRID probe analysis using the 15 compounds of the thumb training set (contour levels: 0.0008 red, −0.0008 blue). To aid in interpretation, only the highest active (**6** in cyan) and one of the lowest active (**11** in orange) compounds are shown. Hydrogen atoms are omitted for the sake of clarity.

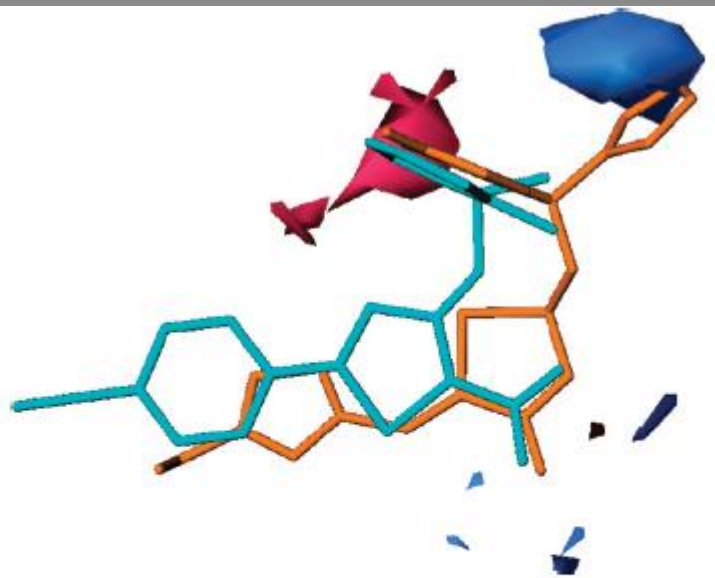


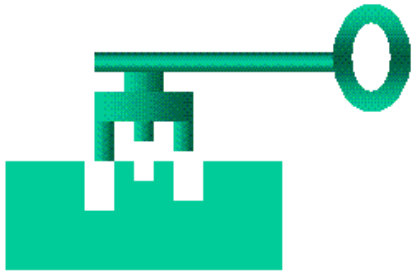
Figure 4. Contour maps of the PLS coefficients derived from OH GRID probe using the 15 compounds of the thumb training set (contour levels: 0.0008 red, −0.0008 blue). To aid in interpretation, only the highest active (**6** in cyan) and one of the lowest active (**11** in orange) compounds are shown. Hydrogen atoms are omitted for the sake of clarity.

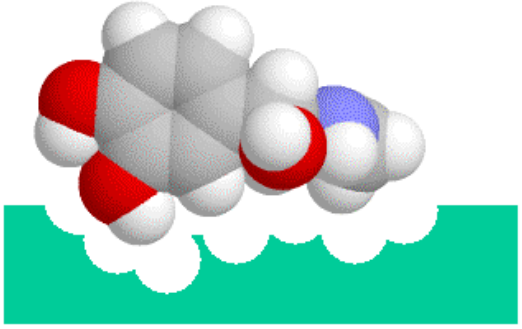
Ricognizione molecolare:

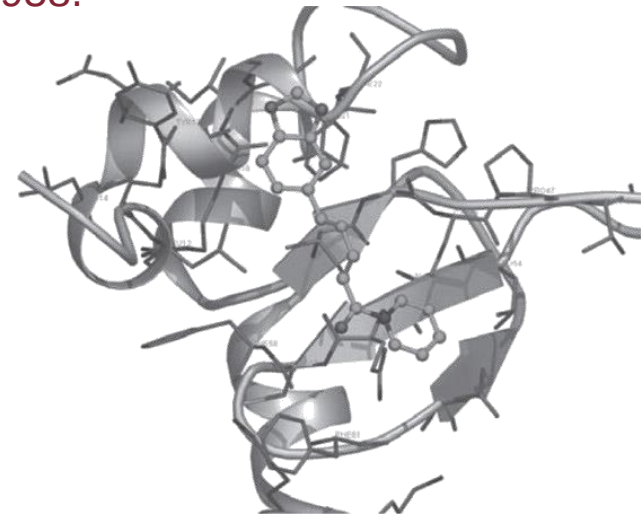
Insieme di interazioni tra molecole e macromolecole

Modello "*Lock* and *key*"
La proteina ha una sua conformazione all'interno del quale il ligando "fitta" perfettamente.
Emil Fischer(1890).

Modello del *fitting* indotto
Il sito attivo della macromolecola può essere modificato a seconda di come interagisce il ligando.
Daniel Koshland 1958.





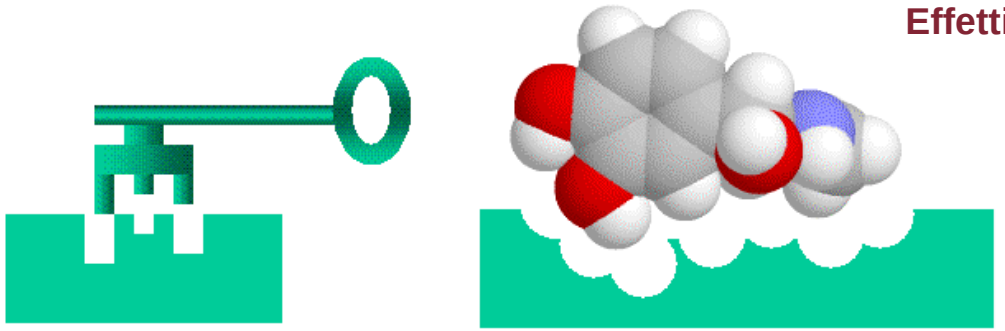


- **Natura delle interazioni?**
- **Intensità della ricognizione molecolare?**

Ricognizione molecolare

La complementarità molecolare altamente specifica tra key (ligando) e lock (recettore) gioca un ruolo chiave nei processi biologici.
La capacità del recettore di agganciarsi al suo ligando con alta specificità e affinità è dovuta alla formazione di una serie di legami deboli e interazioni favorevoli.
Interazioni specifiche ligando-recettore:

Interazioni elettrostatiche (Non-Cov)
Forze di van der Waals (Non-Cov)
Interazioni π - π
Coordinazione con Metalli
Interazioni idrofobiche
Effetti elettromagnetici

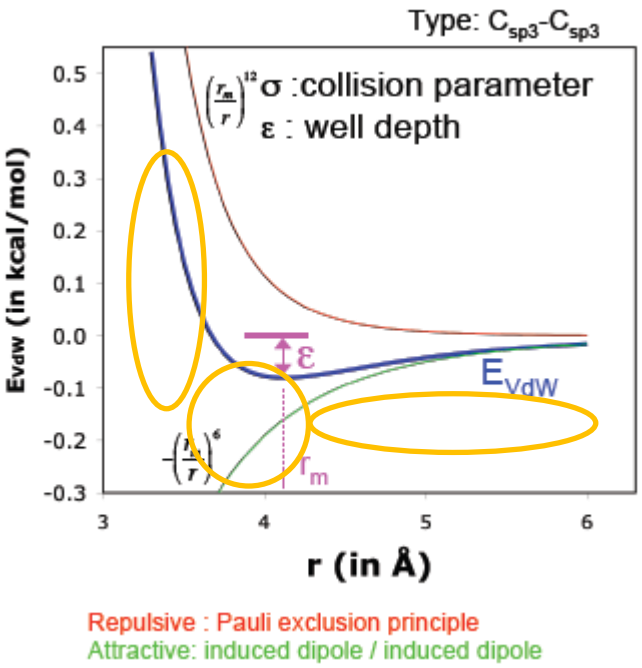


COMPLEMENTARITA' MOLECOLARE

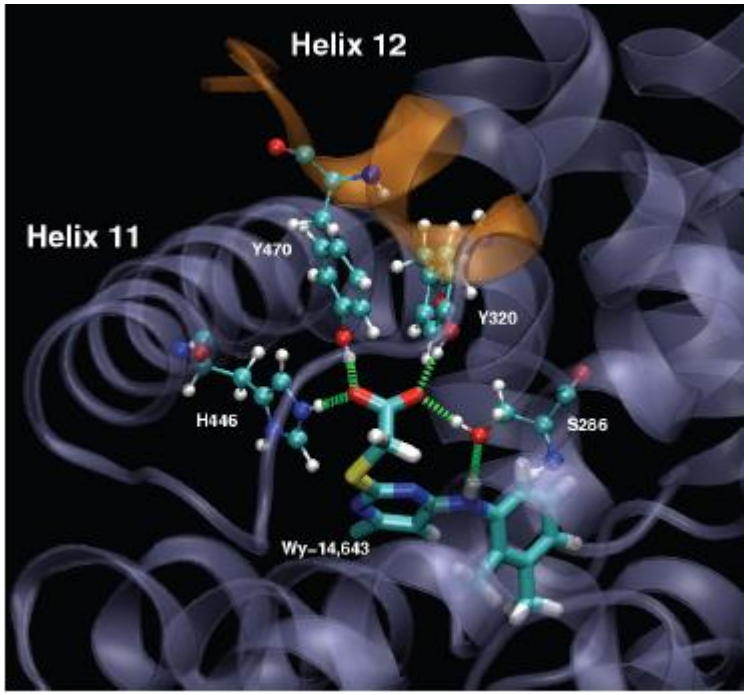
Forze di van der Waals

Debole attrazione intermolecolare causata da dipoli molecolari indotti.

$$E_{VdW} = \varepsilon\left[\left(\frac{r_m}{r}\right)^{12} - 2\left(\frac{r_m}{r}\right)^{6}\right]$$



Repulsive : Pauli exclusion principle
Attractive: induced dipole / induced dipole



Es. Interazione tra il ligando Wy-14,643 e il sito attivo della PPARα

Quando una molecola di ossigeno si avvicina ad una molecola di acqua orientata verso di essa con l'atomo di ossigeno, la frazione di carica negativa presente su quest'ultimo respinge la nuvola elettronica della molecola di ossigeno e attira la carica positiva nucleare. Si determina uno spostamento del baricentro della carica negativa rispetto a quello della carica positiva nella molecola di ossigeno e quindi induce una polarita'.

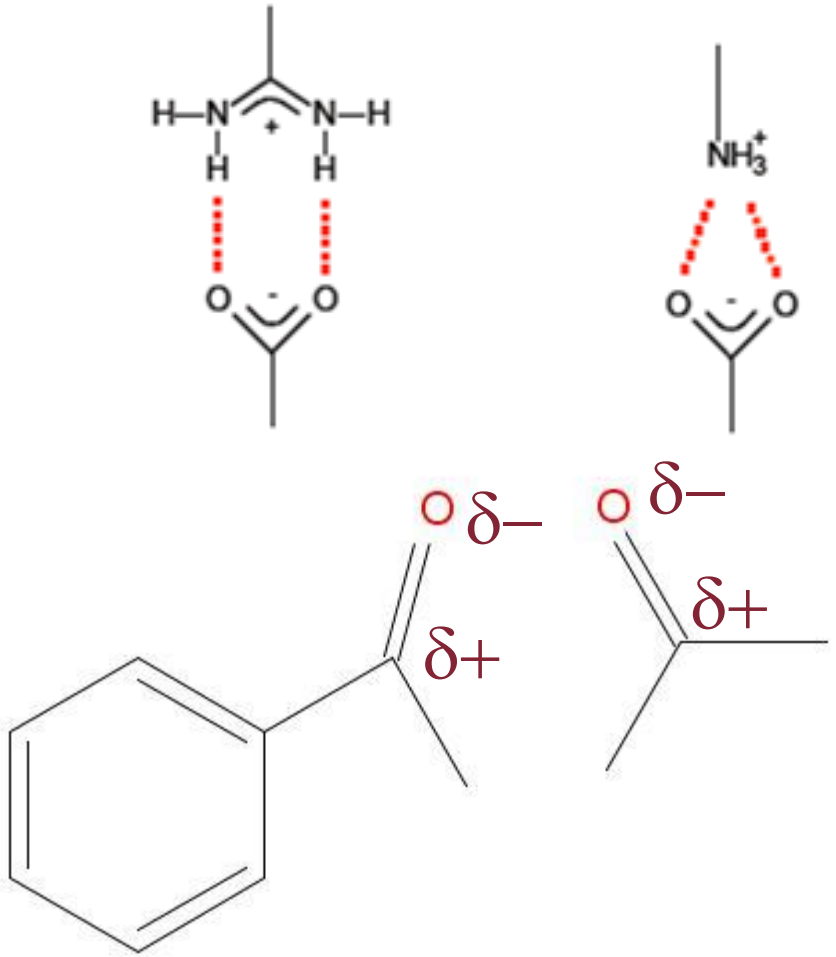
Interazioni Elettrostatiche

La legge di Coulomb

In un mezzo isolante diverso dall'aria la forza *F*, a parità di cariche e di distanza, risulta generalmente minore. La forza di Coulomb si scrive allora:

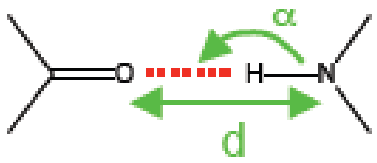
$$\vec{F} = \frac{1}{4\pi \cdot \varepsilon_0 \cdot \varepsilon_r} \cdot \frac{q_1 \cdot q_2}{r^2} \cdot \hat{r}$$

dove **εr** è un numero che prende il nome di costante dielettrica relativa (permittività elettrica relativa). Il prodotto **εr** x **εo** si indica con **ε**, e si chiama costante dielettrica del mezzo (permittività del mezzo).



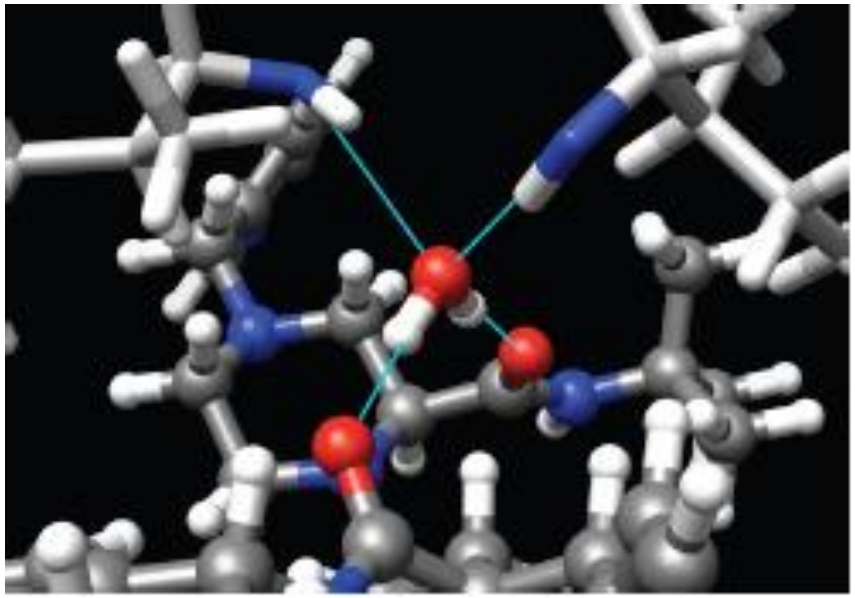
Interazioni Elettrostatiche

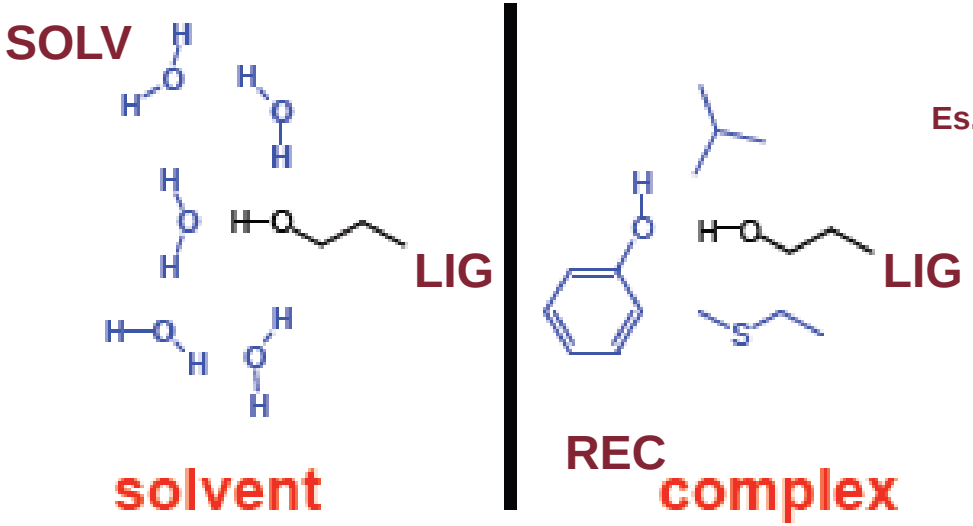
Il legame Idrogeno
ElegH (solv) – ElegH (comp)
determina se i legami-H contribuiscono o no all'affinità



d = 2.5 - 3.2 Å

α = 130 - 180°





Es. Legame-H tra l'Indinavir, una molecola d'acqua e la Ile50 della proteasi dell'HIV-1 (1HSG in PDB)

Il fattore entropico ed entalpico nel binding

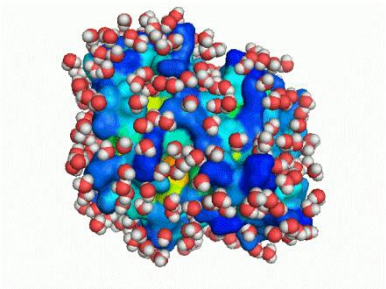
SISTEMA:

La formazione di un complesso è guidata dal cambiamento di **ENTALPIA** ed **ENTROPIA** del sistema
L'equazione fondamentale è:

$$\Delta G = \Delta H - T\Delta S$$

dove **ΔG** = variazione energia libera, **ΔH** = variazione entalpia, **ΔS** = variazione entropia, **T** = temperatura (K)

Effetti di solvatazione (Termine Entropico + Entalpico)



- La ricognizione molecolare tra due molecole avviene in ambiente acquoso
- Le molecole di acqua mediano l'interazione attraverso ponti idrogeno
- Rottura e formazione di legami-H
- La desolvatazione del ligando e della proteina attraverso la complessazione è:

Sfavorevole (Zone Elettrostatiche)

- **Carica del solvente**
- **Screening delle interazioni elettrostatiche sulla superficie della macromolecola**

Favorevole (zone apolari)

- **Cavità nel sito attivo**
- **Riorganizzazione delle molecole d'acqua**
- **Interazioni di tipo van der Waals tra solvente e ligando**
- **Gli effetti di desolvatazione sono proporzionali all'area superficiale accessibile**

Il fattore entropico nel binding

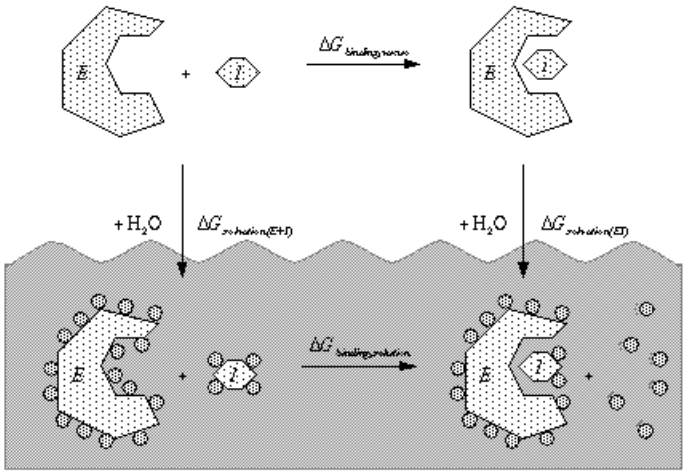
$$\Delta G_{binding,solution} = \Delta G_{binding,vacuo} + \Delta G_{solvation(EI)} - \Delta G_{solvation(E+I)}$$

La **stabilità di un complesso può essere valutata determinando la costante di equilibrio, che è correlata alla variazione di energia libera:**

$\Delta G = \Delta G0 - RT \ln Kd$

dove ΔG0=variazione dell'energia libera in condizioni standard, R= costante dei gas,T=temperatura assoluta, Kd = costante di legame espressa come costante di dissociazione con

[R] + [L] ↔ [RL]



reazione di formazione del complesso è:

Kd = [R] [L] / [RL]

Kd è definita come la concentrazione del ligando per la quale il 50% dei siti del recettore sono occupati.

L'**effetto idrofobico** è il termine maggiormente stabilizzante per i complessi biomolecolari, mentre le **interazioni coulombiane** e i **legami idrogeno** forniscono specificità alle interazioni proteina-ligando.

Il fattore entropico nel binding

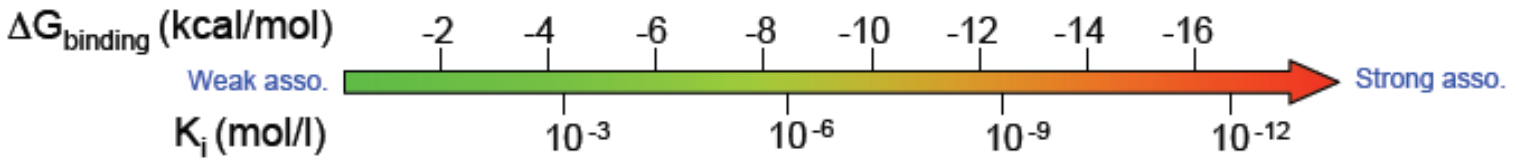
Binding, dissociation and inhibition constants.
Binding free energy

$$A + B \leftrightarrows A'B' \qquad K_b = \frac{[A'B']}{[A][B]}$$

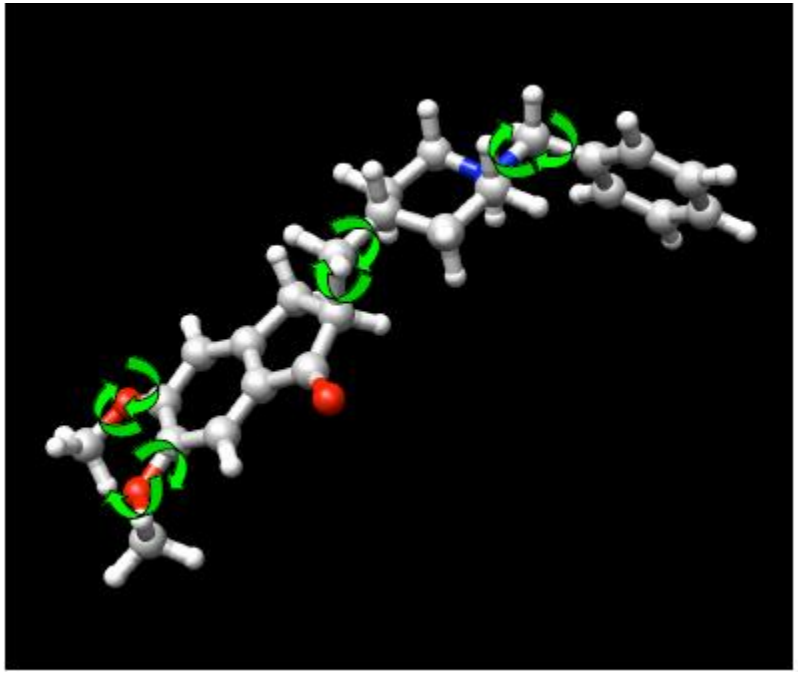
$$A'B' \leftrightarrows A + B \qquad K_d = K_i = \frac{[A][B]}{[A'B']}$$

K_b : binding constant, K_d : dissociation constant, K_i : inhibition constant

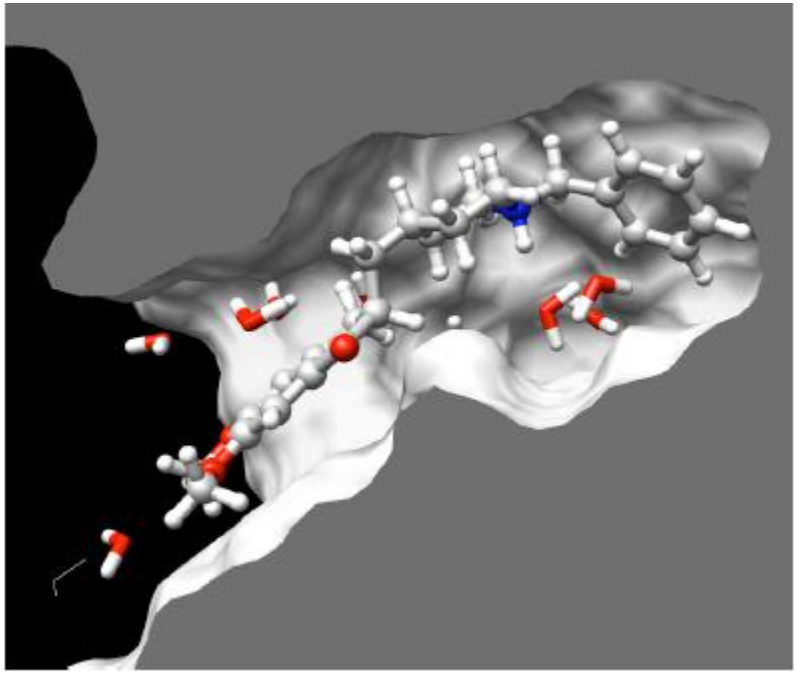
$$\Delta G_{binding} = -RT\ln K_b = RT\ln K_i = \Delta H - T\Delta S$$



Il fattore entropico nel binding



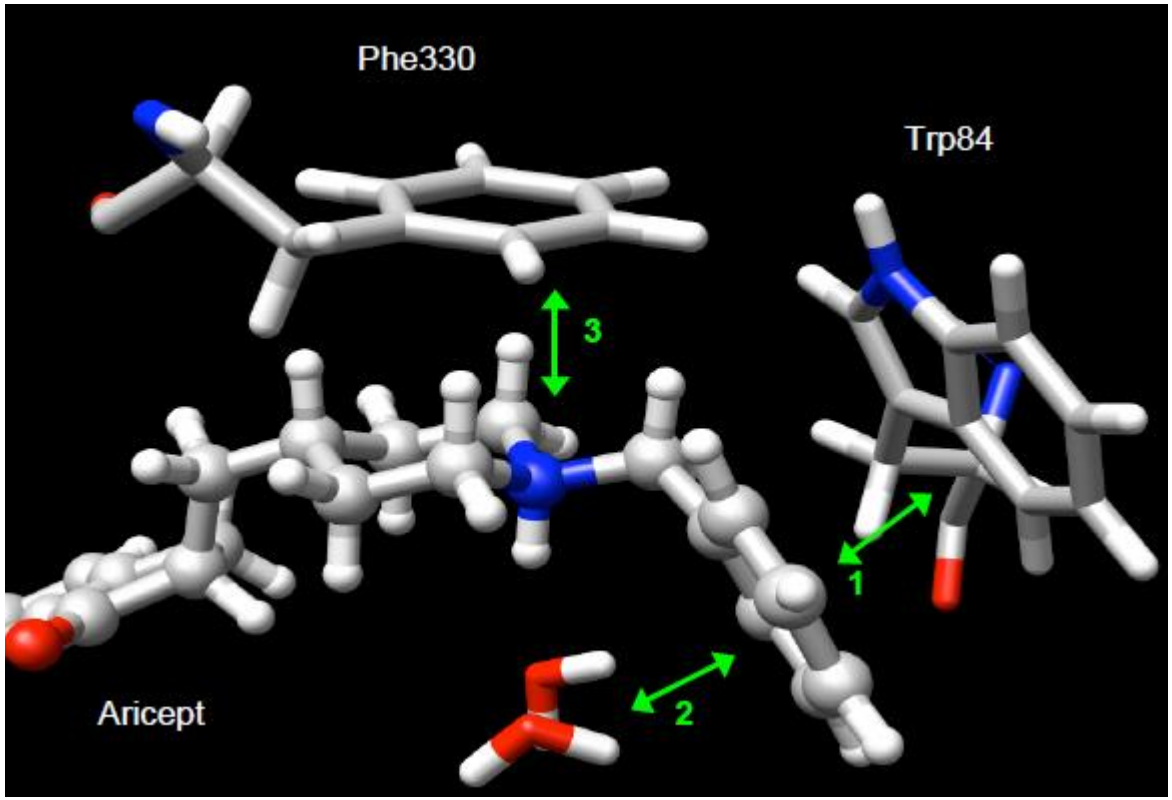
Es. Aricept in soluzione



Es. Aricept nel sito attivo
della acetilcolinesterasi

I gradi di libertà “congelati” durante la complessazione rendono il ligando sfavorito al *binding*

Interazioni π

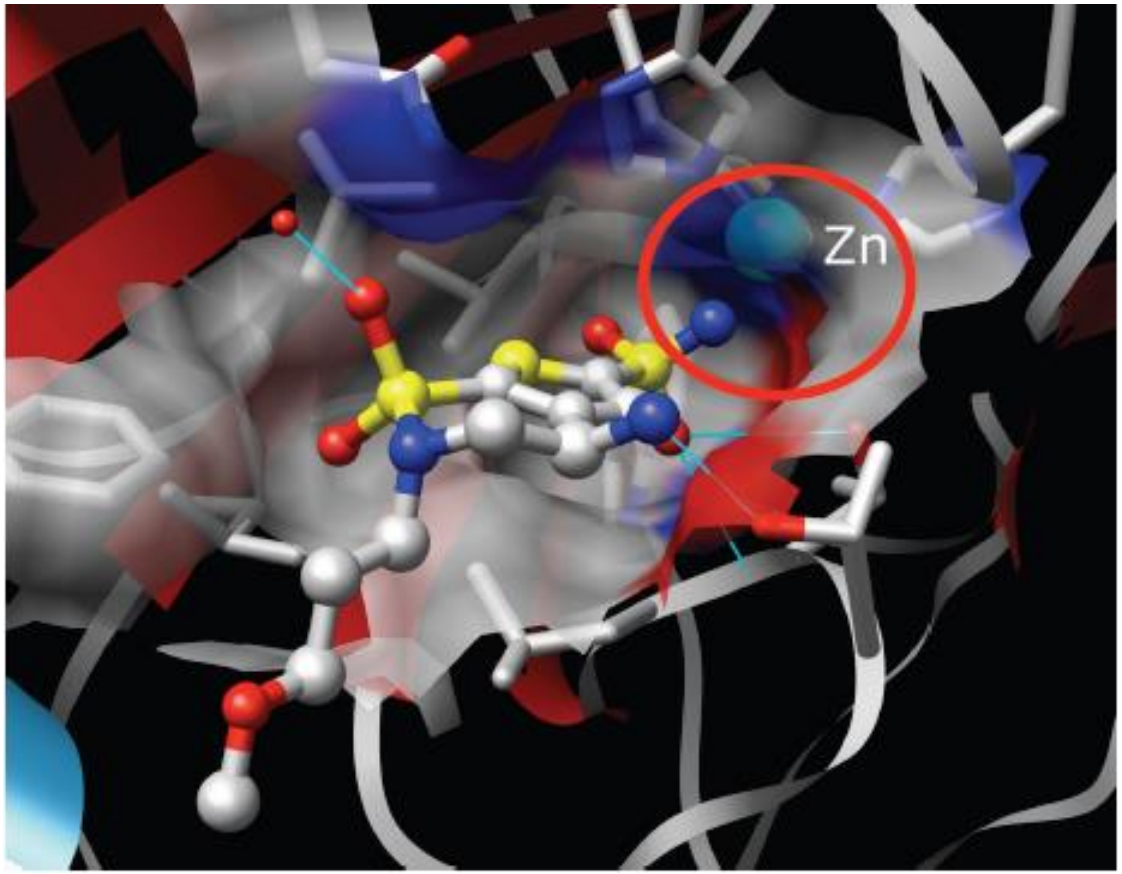


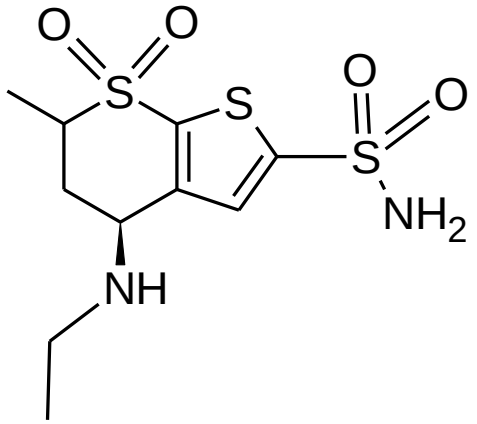
Es. Interazioni π tra l'Aricept (farmaco anti-Alzheimer approvato dalla FDA) e l'acetilcolinesterasi

1. Interazioni π-π
2. Interazioni OH-π
3. Interazioni cationi-π

Coordinazione con Metalli

Legame covalente di coordinazione





Dorzolamide: Inibitore dell'anidrasi carbonica, approvato dalla FDA come
agente anti-glaucoma

DOCKING: Le *Scoring Functions*

Structure-based With 3D structure for the targeted macromolecule	- Physical based scoring functions : • Free energy simulations (FEP, TI) • MM-PBSA, MM-GBSA • Linear interaction energy (LIE) - Empirical-based scoring function (regression-based approaches) Ex: Ludi score - Knowledge-based approaches (Potentials of Mean Force, PMF) Ex: PMF score

Utility?

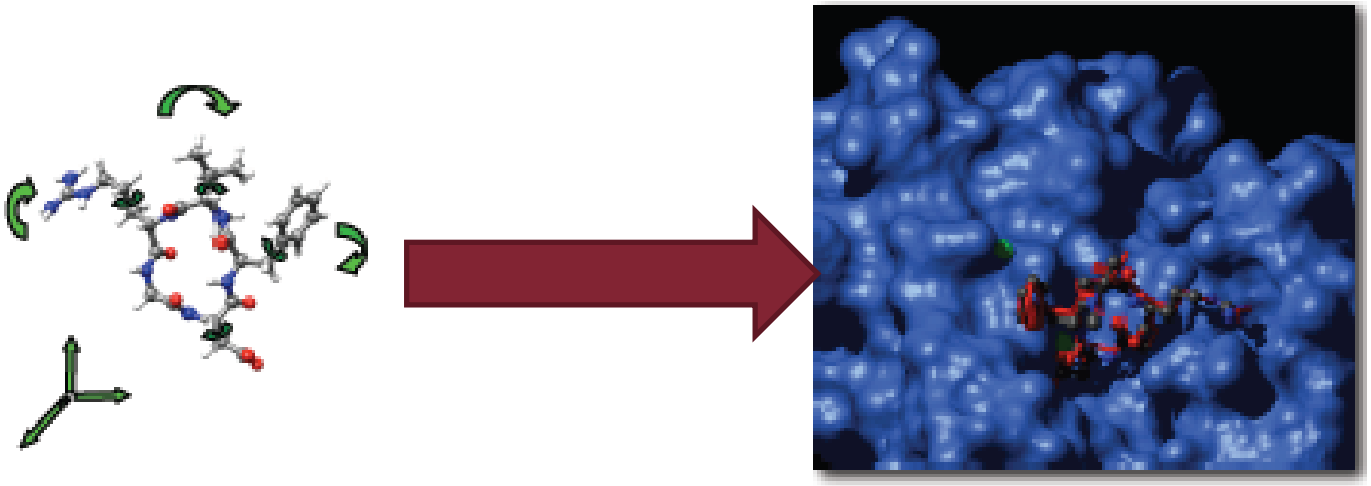
- *Virtual Screening*
- *De novo design*
- *Lead optimization*

Ligand Docking

Posa o ***Binding Mode*:** posizione, orientamento e conformazione di un ligando sulla superficie di macromolecole biologiche

***Binding Mode* sperimentale:** Estratta da dati cristallografici e stimata come la migliore posa in termini di *free binding energy.*

Docking: Metodo computazionale utilizzato per la predizione di un binding mode il più vicino possibile allo sperimentale (<2 Å)



Docking Software

Program	ALGORITHM
AutoDock	Lamarckian GA
DOCK	Shape matching (sphere images)
DOCK (NWU version)	Shape matching (sphere images)
FlexX	Incremental construction
FRED	Shape matching (gaussian functions)
Glide	Descriptor matching/MC
GOLD	GA
Hammerhead	Incremental construction
ICM	MC minimization
LigandFit	Shape matching (moments of inertia)
QXP	MC minimization, tree searching and pruning
SLIDE	Descriptor matching
Surflex Dock	Surface-based molecular similarity

Classifying Docking Algorithms

- Ligand Conformations
  - Rigid (i.e. protein-protein in bioinformatics)
  - Fixed Sample (i.e. ligand-protein in medicinal chemistry)
  - Flexible (i.e. molecular dynamic in thermodynamic calcs)
- Constraints
  - Residues and Pockets
  - Pharmacophores
- Charges and Tautomers
- Water Molecules

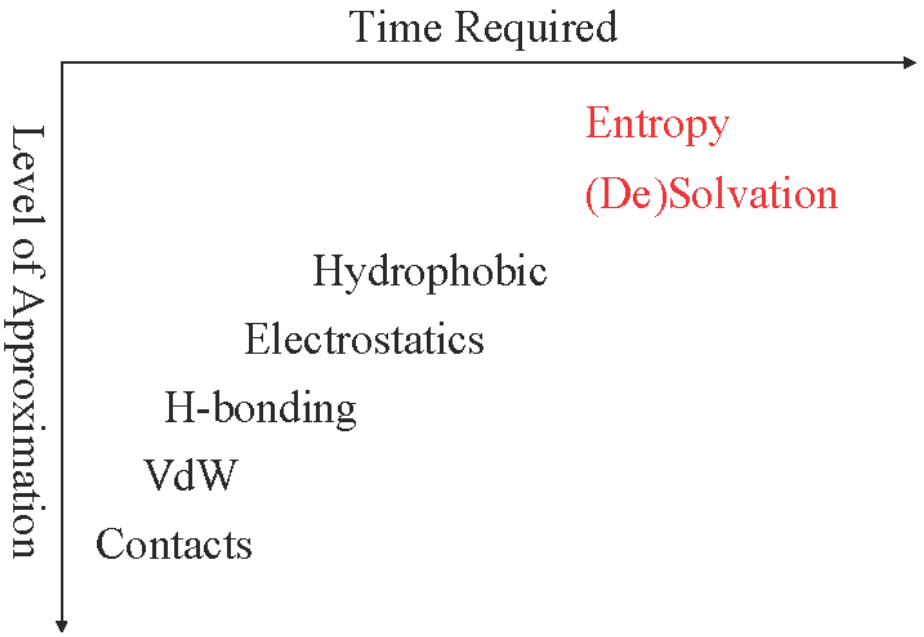
Water Molecules

- Essential
  - Mediates Binding for all Ligands
- Optional
  - Presence Required by Some Ligands
  - Inhibits Binding of Other Ligands
- Solvation & Desolvation Free Energy Critical for Scoring Function

Limitation for Current Methods

- Rigid Proteins Side-Chains
- Binding Relevance for Biological Activity
- Scoring Functions

Terms in Scoring Functions



Classifying Scoring Functions

- Knowledge Based (Atoms Pairs in Contact)
  - DrugScore, PMF
- Energ
  - GOLD, DOCK, LigandFit, MOE
- Energy + Parametrized Solvations
  - Chemscore (Glide, THINK)
- Free Energy Perturbation

Knowledge-Based Functions

$$\text{Score} = \Sigma_{r<\text{cutoff}} A_{ij}(r)$$

- Potentials of Mean Force (PMF)
  - J Med Chem (1999) 42 p791-804
- DrugScore
  - J Med Chem (2005) 48 p6296-6303
- Less Confused by Crystal Structure Precision

Energy

- Lennard Jones

  $\Sigma_{ij} (A / r_{ij}^{12} - B/r_{ij}^{6})$

- Torsion Term

  $\Sigma (1 - \cos 2\omega)$ Conjugated

  $\Sigma (1 + \cos 3\omega)$ Non-conjugated

- Electrostatics

  $\Sigma_{ij} q_i q_j / \varepsilon r_{ij}$

Enhanced ChemScore

$\Delta G = \Delta G_0 + \Delta G_{hbond} * N_{hbond} + \Delta G_{lipo} * N_{lipo} + \Delta G_{bad} * N_{bad} + \Delta G_{rot} * N_{rot} + E$

where

ΔG_0 ΔG_{hbond} ΔG_{lipo} ΔG_{bad} ΔG_{rot} are constants (-5.48;-3.34; -0.117; 0.058; 2.56)

N_{hbond} is the number of interactions (using geometric criteria)

N_{lipo} is the number of lipohilic contacts (cf PMF, DrugScore)

N_{bad} is the number of lipohilic-hydrophilic contacts (extension)

N_{rot} is the number of frozen rotatable bonds in the ligand

E is the VdW interaction energy and ligand torsional energy (extension)

Free Energy Perturbation

$$\Delta G_{sol} = -\Delta G_{sol}(\text{ligand}) - \Delta G_{sol}(\text{protein}) + \Delta G_{gas} + \Delta G_{sol}(\text{complex})$$

- Error Prone due to Subtraction of Large Numbers
- Solvent Accessible Surface Area (SASA) approximation (cf ChemScore)
- J Med Chem (2004) 47 p3065-74

Validation and Results

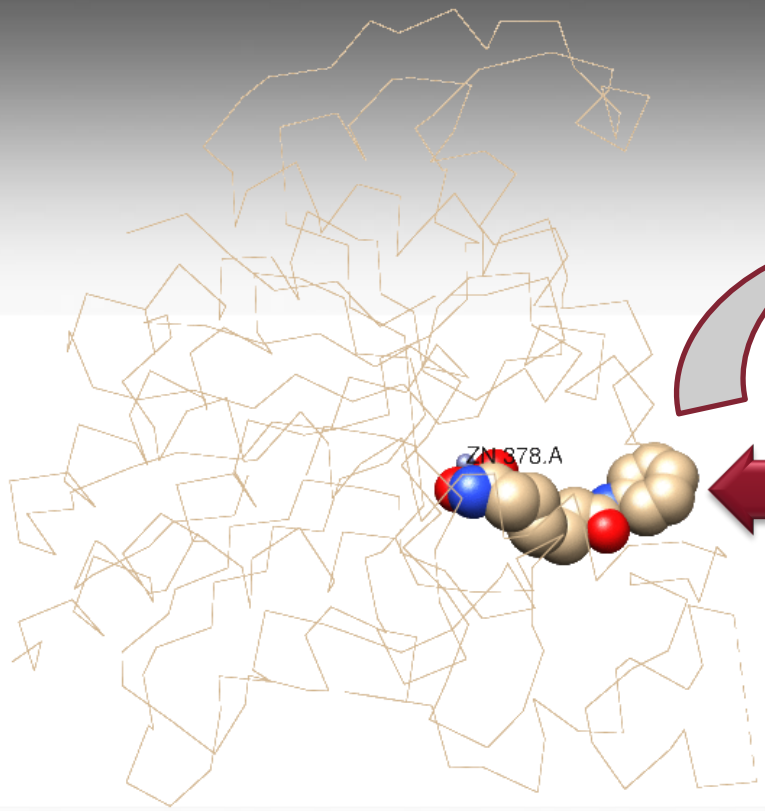
- Reproduce Ligand-Protein Crystal Structures
  - RMS Deviation of non-H Atoms
  - Docking Score
- Dock Actives
  - Used for Developing Scoring Functions
- Prediction
  - Enrichment over Random
  - Percentage of False Positives

What We Can and Cannot Do

- Routine
  - Small molecule conformation generation and energy profiling
  - Visualizing crystal structures
  - Binding site characterization
  - Virtual screening to enrich databases for actives
    - Cheminformatics, ligand-based, and structure-based
  - Predict binding modes when receptor can be treated rigidly
- Difficult
  - Separating highly from weakly active compounds
  - Predicting side chain rearrangements and backbone relaxation
- Very Challenging
  - Predicting binding free energies
  - Predicting large scale protein movements
  - Mapping free energy surfaces
  - Understanding off-target effects
  - Other ADME-Tox

SCHRÖDINGER.

COMBINE



SAHA

Journal of
Medicinal Chemistry

Prediction of Drug Binding Affinities by Comparative Binding Energy Analysis

Angel R. Ortiz, M. Teresa Pisabarro, Federico Gago, and Rebecca C. Wade

J. Med. Chem., 1995, 38 (14), 2681-2691• DOI: 10.1021/jm00014a020 • Publication Date (Web): 01 May 2002

	Res1	Res2	Res3	.....	Res-n
Cplx 1	X1	X2	X3	.....	Xn
Cplx 2	X1	X2	X3	.....	Xn
Cplx 3	X1	X2	X3	.....	Xn
...	.....				
...	.....				
...	.....				
...	.....				
...	.....				
...	.....				
...	.....				
...	.....				
...	.....				
Cplx m	X1	X2	X3	.....	Xn

PLS

Y1
Y2
Y3
...
...
...
...
...
...
...
...
...
Ym

$$\Delta U = \sum_{i=1}^{n_l} \sum_{j=1}^{n_r} u_{ij}^{\mathrm{VDW}} + \sum_{i=1}^{n_l} \sum_{j=1}^{n_r} u_{ij}^{\mathrm{ELE}}$$

COMBINE Model

COMBINE for Selectivity

PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Comparative binding energy analysis for binding affinity and target selectivity prediction

Stefan Henrich,[1] Isabella Feierberg,[2] Ting Wang,[1] Niklas Blomberg,[2] and Rebecca C. Wade[1]*

Proteins: Structure, Function, and Bioinformatics, 2010, 78, 135–153

Methods

COMBINEr
COMBINE with Autogrid/R

www.rcmd.it

JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

3-D QSAutogrid/R: An Alternative Procedure To Build 3-D QSAR Models. Methodologies and Applications

Flavio Ballante[†] and Rino Ragno*[,†]

J Comput Aided Mol Des
DOI 10.1007/s10822-012-9586-6

Comprehensive model of wild-type and mutant HIV-1 reverse transciptases

Flavio Ballante · Ira Musmuca · Garland R. Marshall · Rino Ragno

Journal of Medicinal Chemistry

Brief Article

pubs.acs.org/jmc

2-(Alkyl/Aryl)Amino-6-Benzylpyrimidin-4(3*H*)-ones as Inhibitors of Wild-Type and Mutant HIV-1: Enantioselectivity Studies

Dante Rotili,[†,#] Alberta Samuele,[‡,#] Domenico Tarantino,[†] Rino Ragno,[†] Ira Musmuca,[†] Flavio Ballante,[†] Giorgia Botta,[†] Ludovica Morera,[†] Marco Pierini,[†] Roberto Cirilli,[§] Maxim B. Nawrozkij,[∥] Emmanuel Gonzalez,[⊥] Bonaventura Clotet,[⊥] Marino Artico,[†] José A. Esté,*[,⊥] Giovanni Maga,*[,‡] and Antonello Mai*[,†]

ACS Publications   dx.doi.org/10.1021/jm201308v | J. Med. Chem. 2012, 55, 3558–3562

Methods

COMBINEr
COMBINE with Autogrid/R

JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

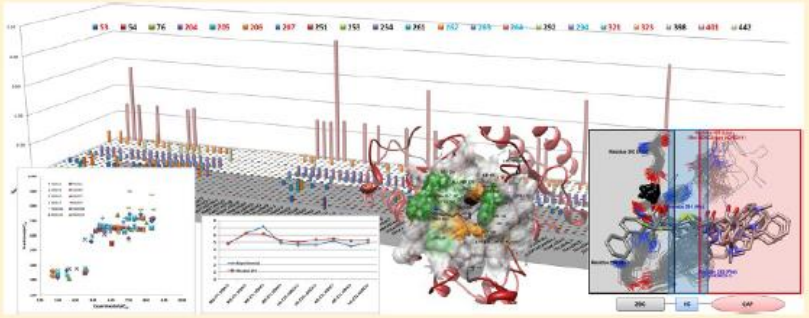
pubs.acs.org/jcim

Histone Deacetylase Inhibitors: Structure-Based Modeling and Isoform-Selectivity Prediction

Laura Silvestri,[†] Flavio Ballante,[†] Antonello Mai,[‡] Garland R. Marshall,[†,§] and Rino Ragno*[,†]

[†]Rome Center for Molecular Design Dipartimento Chimica e Tecnologie del Farmaco, Facoltà di Farmacia e Medicina, [‡]Istituto Pasteur—Fondazione Cenci Bolognetti Dipartimento Chimica e Tecnologie del Farmaco, Facoltà di Farmacia e Medicina, Sapienza Università di Roma, Rome, Italy

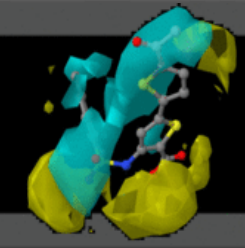
Supporting Information



ABSTRACT: An enhanced version of comparative binding energy (COMBINE) analysis, named COMBINEr, based on both ligand-based and structure-based alignments has been used to build several 3-D QSAR models for the eleven human zinc-based histone deacetylases (HDACs). When faced with an abundance of data from diverse structure−activity sources, choosing the best paradigm for an integrative analysis is difficult. A common example from studies on enzyme−inhibitors is the abundance of crystal structures characterized by diverse ligands complexed with different enzyme isoforms. A novel comprehensive tool for data mining on such inhomogeneous set of structure−activity data was developed based on the original approach of Ortiz, Gago, and Wade, and applied to predict HDAC inhibitors' isoform selectivity. The COMBINEr approach (apart from the AMBER programs) has been developed to use only software freely available to academics.

Methods

HDAC/COMBINEr Protocol



RCSB PDBProtein Data Bank

Web Servers:
- CPHmodels 3.0,
- ModWeb,
- SwissModel,
- M4T 3.0.

Experimental Structures

Homology Modeling

HDAC -2,-4,-7,-8

HDAC -1,-3,-5,-6, -9,-10,-11

Multiple Alignment (Modeller)

Training Set

Test Sets

Training Set

Test Sets

Complexes Energy Minimization (Amber 10)

Molecular Docking (AutoDock Vina)

Training Set 94 complexes HDAC-inhibitors

Test Set 76 complexes HDAC-inhibitors

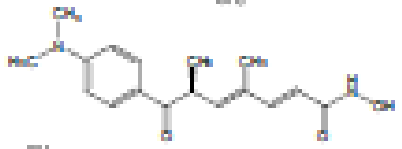
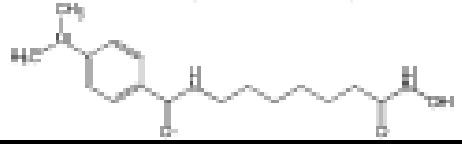
Calculation of Molecular Interaction Energies on a per Residue Basis

Statistical Analysis Partial Least Square Regression (PLS):
- Model,
- Cross Validation,
- Y-Scrambling
COMBINE Models

Predictability Assessment by Three External Test Sets

Methods: Workflow

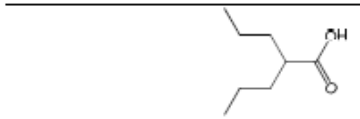
COMBINEr Training Set

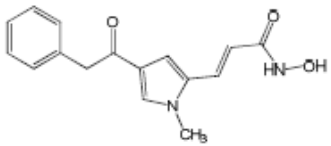
PDB code (ligand name)	Ligand structure	PDB code (ligand name)	Ligand structure
3MAX (LLX)		1T69 3C0Z (SAHA)	
3F07 (APHA8)		1W22 (NHB)	
1T64 3C10 (TSA)		2VQM (HA3)	
1T67 (MS-344)		2VQJ (TFMK)	

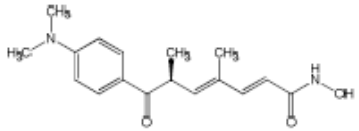
PDB code	HDAC Class	HDAC Number	IC_{50} (μM)	PDB code	HDAC Class	HDAC Number	IC_{50} (μM)
3MAX[11] (LLX)	I	2	0.9	1W22[14] (NHB)	I	8	0.175
3F07[12] (APHA8)	I	8	2.8	2VQM[17] (HA3)	II a	4	0.978
1T64[14] (TSA)	I	8	1.1	2VQJ[17] (TFMK)	II a	4	0.367
1T67[14] (MS-344)	I	8	0.249	3C0Z[18] (SAHA)	II a	7	0.05
1T69[14]	I	8	2.2	3C10[18] (TSA)	II a	7	0.014

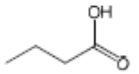
Experimental Bound HDACi

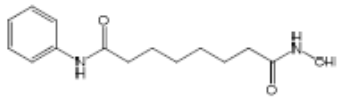
COMBINEr Training Set

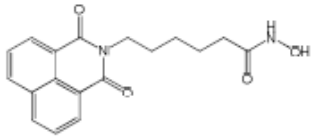

valproic acid (VALP)

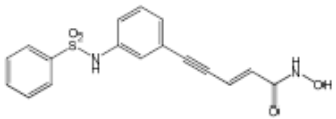

APHA8

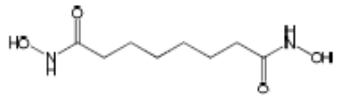

TSA

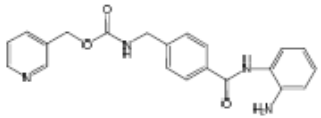

Butyrate (NABUT)


SAHA


SCRIPTAID (SCRIP)


Oxamflatin (OXAM)


SBHA


MS-275

HDAC	CLASS	I				IIa				IIb		IV
	Number	1	2	3	8	4	5	7	9	6	10	11
Chemical Structures and IDs	valproic acid (VALP)	1000	1000	226.08	228.85	-	2000	-	2000[19]	1000	1000	-
	Butyrate (NABUT)	319	28.9	22.5	85.6	30	2000	30	2000[19]	1000	292	-
	Oxamflatin (OXAM)	0.05	0.2	0.01	2.2	0.03	-	0.03	-	0.09	0.05	-
	APHA8	3.7	7.4	0.42	2.8	3.1	-	3.1	-	0.1	4.2	-
	SAHA	0.1	0.44	0.02	2.2	0.05	0.378	0.05	0.316	0.02	0.1	0.362
	SBHA	2.1	4.6	0.41	3.7	1.4	-	1.3	-	0.1	2.3	-
	TSA	0.005	0.021	0.005	1.1	0.014	0.0165	0.014	0.0381	0.005	0.005	0.0152
	SCRIPTAID (SCRIP)	0.17	0.64	0.03	2.3	0.2	-	0.16	-	0.004	0.17	-
	MS-275	13	0.51	0.07	30	12	-	6.2	-	21	11.5	-

Life Sciences 82 (2008) 1050–1058

Contents lists available at ScienceDirect

Life Sciences

journal homepage: www.elsevier.com/locate/lifescie

The use of diversity profiling to characterize chemical modulators of the histone deacetylases

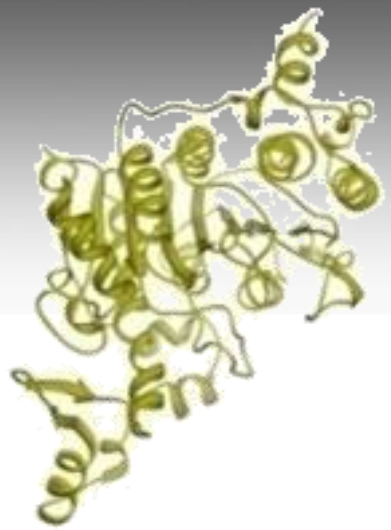
Leonard Blackwell, Jacqueline Norris, Carla M. Suto, William P. Janzen *

Modeled HDACi

Experimental Data

HDAC2

HDAC4

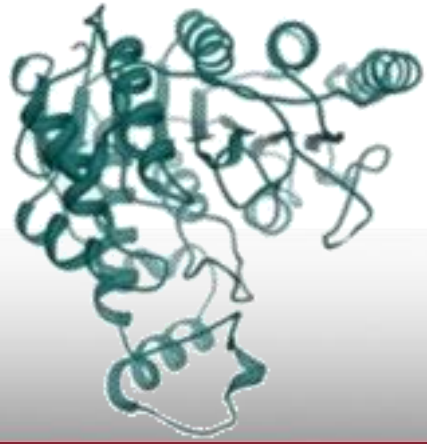


RCSB PDB PROTEIN DATA BANK

Experimental Structures

HDAC7

HDAC8



Homology Modeling

CPHmodels 3.0 Server

ModWeb Server

SwissModel Server

M4T Server 3.0

4 Homology Models

For Each

HDAC-1,-3,-5,-6A,-6B,-9,-10,-11

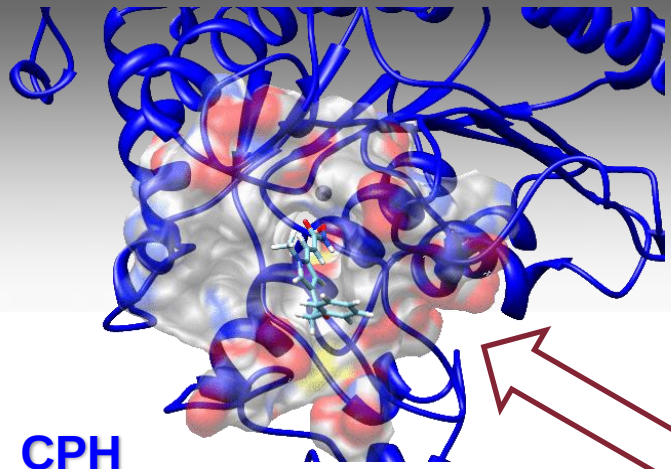
Docking Assessment

$$DA\ (Docking\ Accuracy) = f_{RMSD \leq 2} + 0.5\ (f_{RMSD \leq 3} - f_{RMSD \leq 2})$$

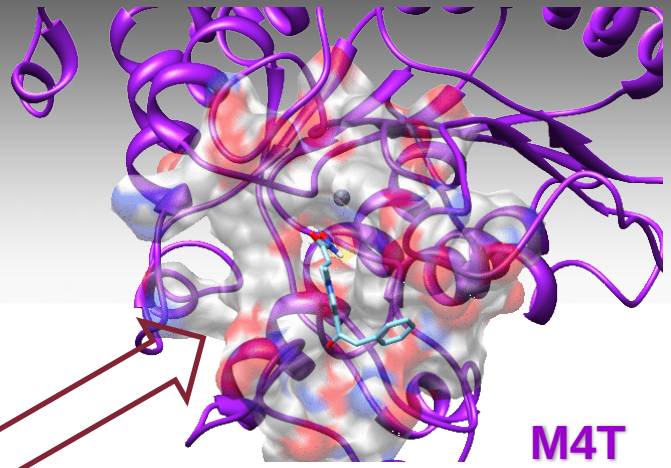
Redocking results (RMSD) with AutoDockVina

Complex name	Best docked	Best Cluster	Best Fit
LLX.HDAC2	0.24	0.24	0.24
HA3.HDAC4	3.87	2.34	1.93
TMFK.HDAC4	4.02	1.9	1.46
SAHA.HDAC7	2.45	2.45	1.88
TSA.HDAC7	2.19	2.19	1.21
APHA.HDAC8	1.43	1.43	1.43
SAHA.HDAC8	2.49	2.49	1.72
TSA.HDAC8	2.09	1.22	1.22
DA %	50	75	100

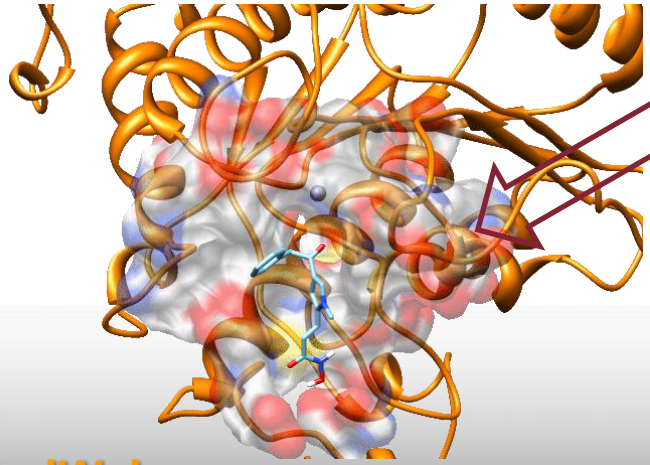
Docking Assessment



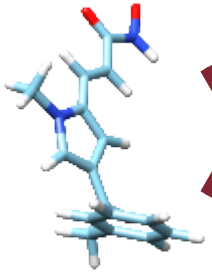
CPH

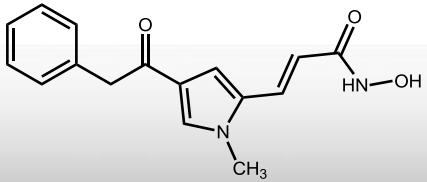


M4T

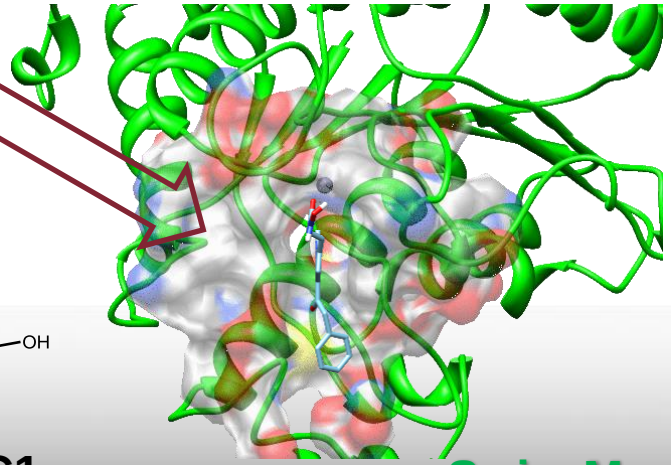


ModWeb



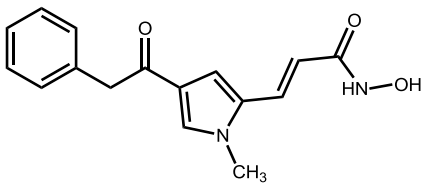


APHA8 in HDAC1

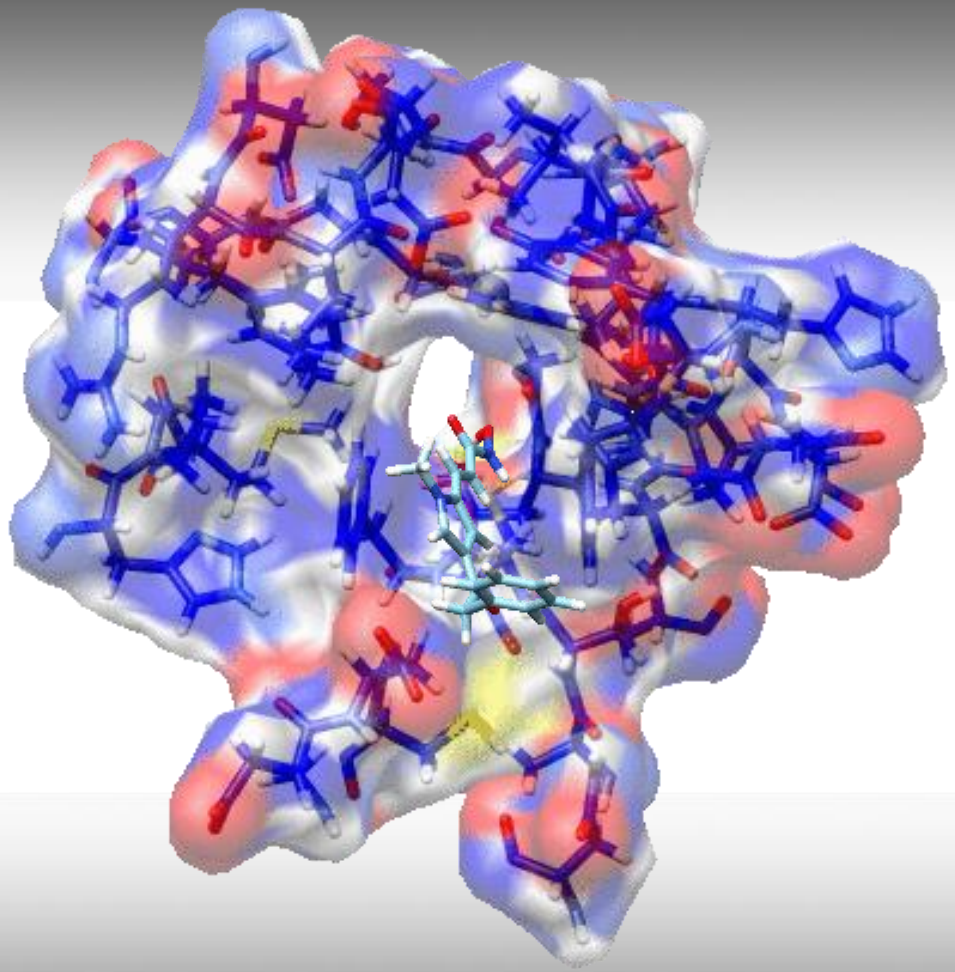


SwissModel

Cross-Docking → Flex-Docking



APHA8 in HDAC1



CPH

ModWeb

M4T

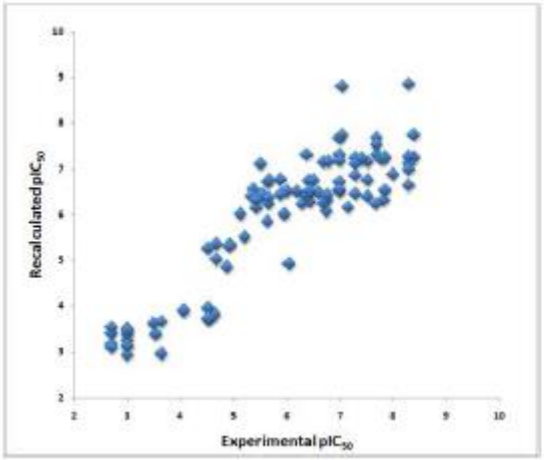
SwissModel

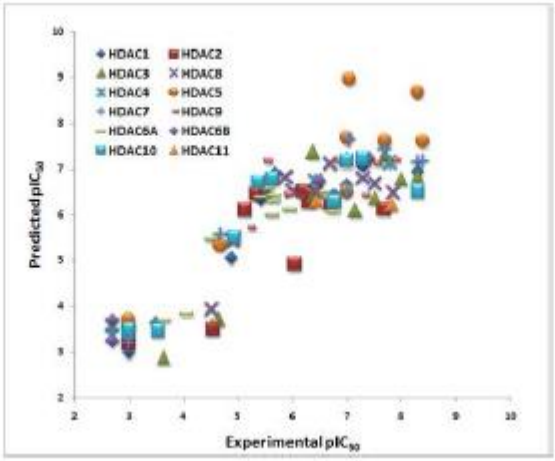
Structure-Based Approach

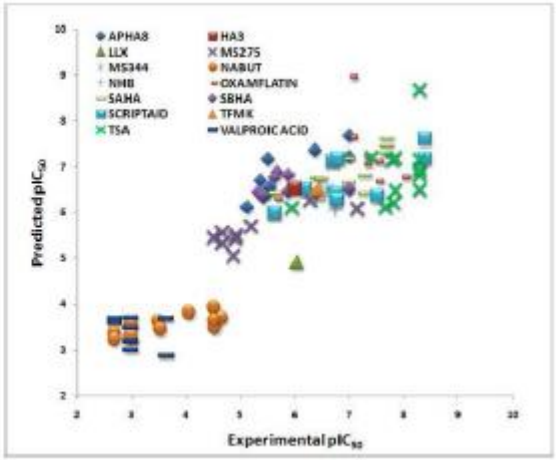
COMBINEr Model Building

Statistical results of the COMBINEr models.

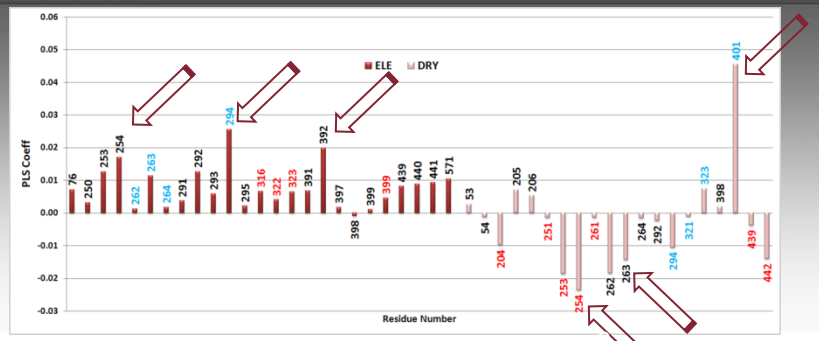
#	Field	PC	r^2	SDEC^c	q^2_{K5fold}	SDEP $_{R5G-LSO}$	q^2_{LOO}	SDEP$_{LOO}$	scrambled q^2 % positive values	scrambled q^2 Max. value
1	ELE	2	0.69	0.91	0.67	0.94	0.68	0.93	5	0.07
2	STE	2	0.27	1.40	0.14	1.52	0.15	1.51	n.d.	n.d.
3	DRY	2	0.46	1.21	0.34	1.33	0.36	1.32	n.d.	n.d.
4	ELE+STE	2	0.74	0.84	0.68	0.93	0.68	0.93	2	0.05
5	ELE+DRY	2	0.80	0.73	0.76	0.81	0.76	0.81	6	0.08
6	STE+DRY	3	0.54	1.11	0.33	1.34	0.35	1.33	n.d.	n.d.
7	ELE+DRY+STE	2	0.77	0.78	0.72	0.87	0.72	0.87	4	0.04

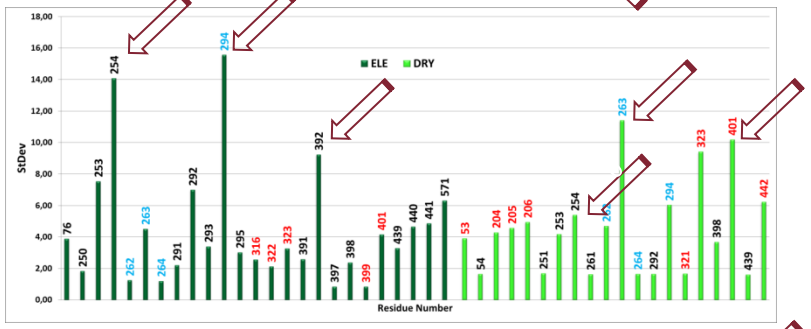


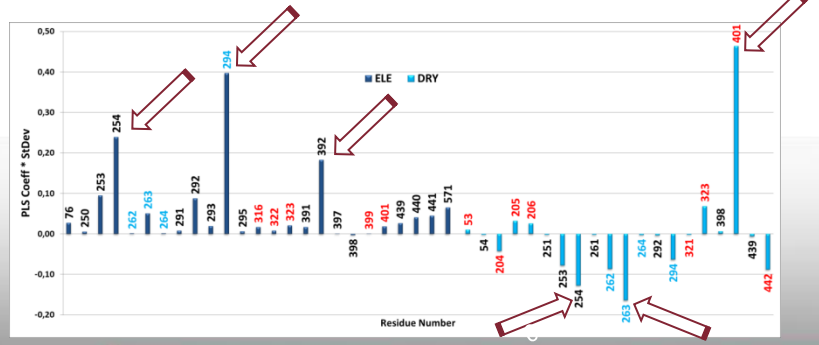


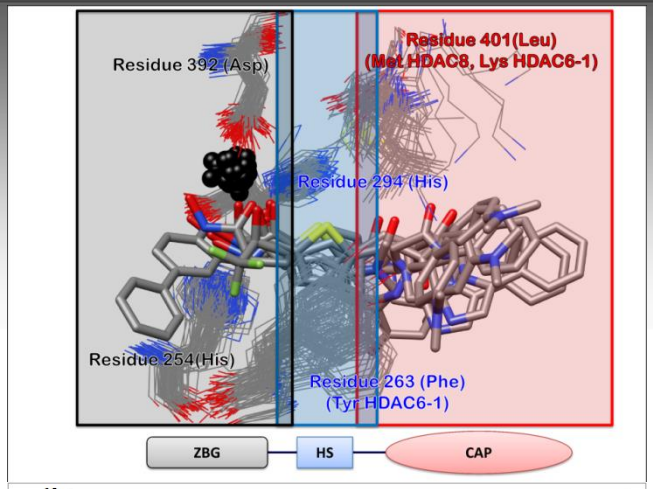


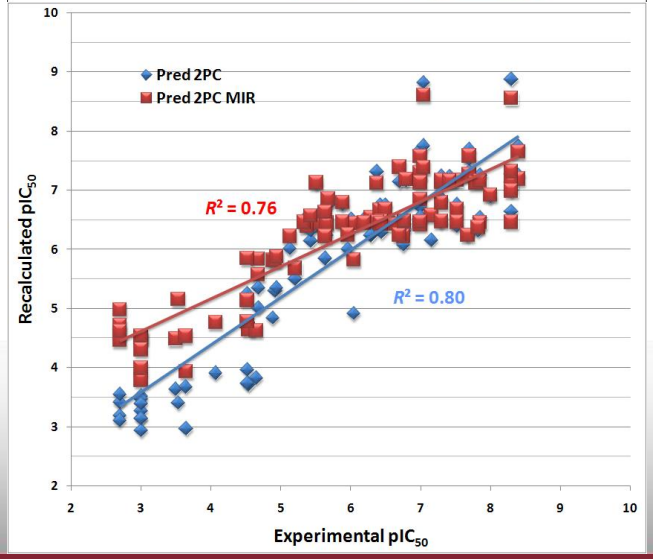
COMBINEr Model Interpretation







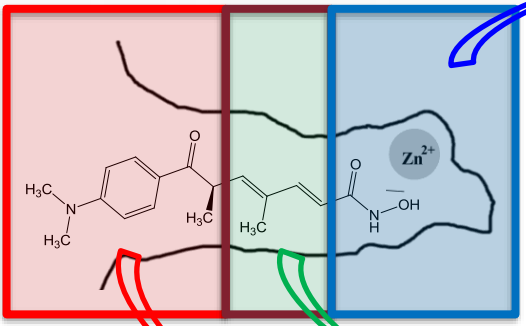


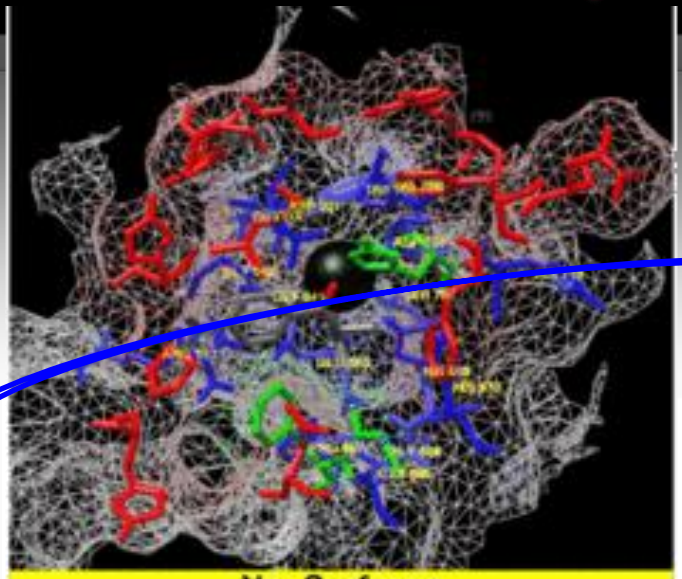


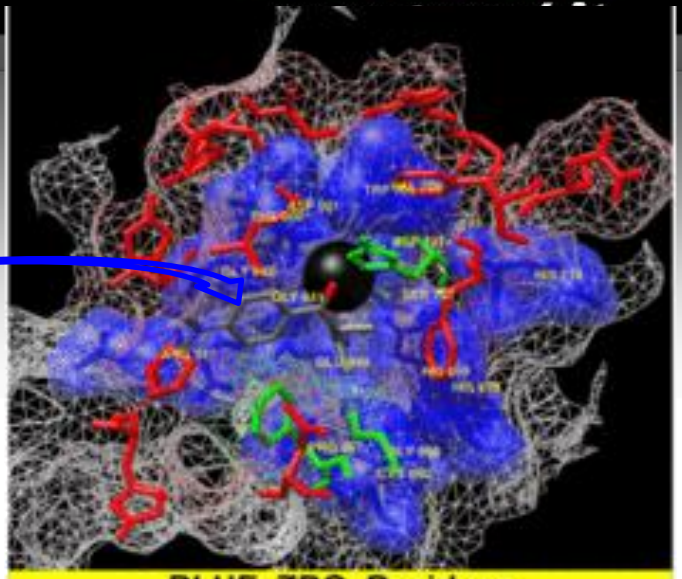
HDAC Inhibitors: Structure-Based Modeling and Isoform-Selectivity Prediction

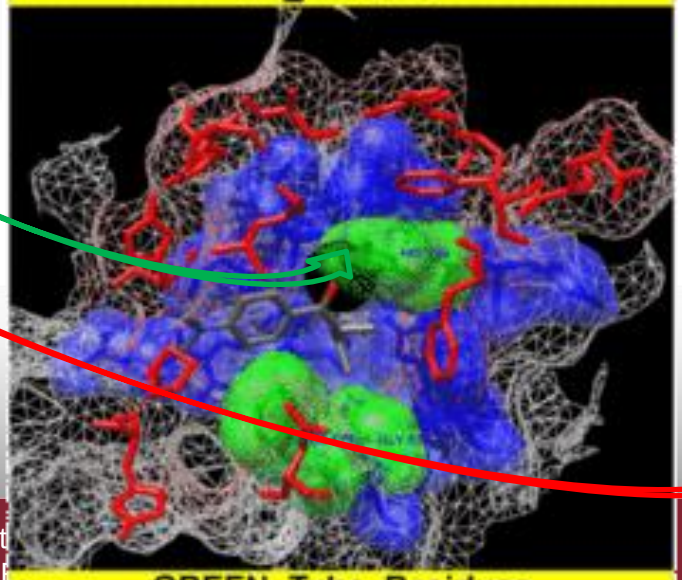
SB 3-D QSAR

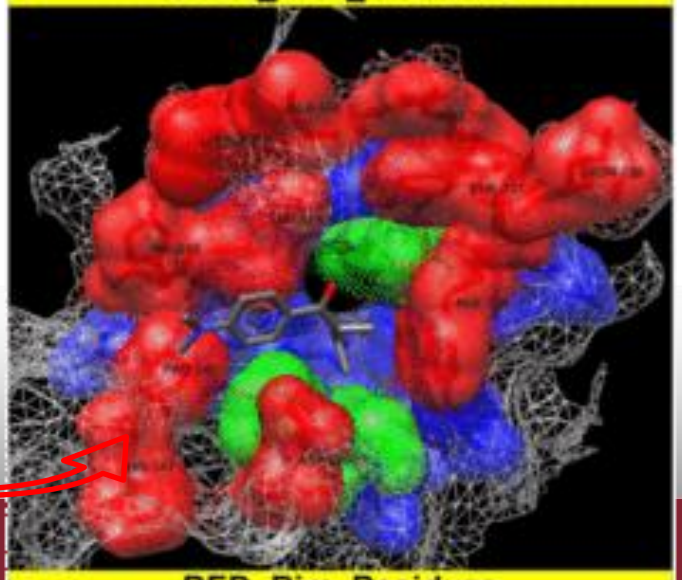
COMBINEr Model Interpretation





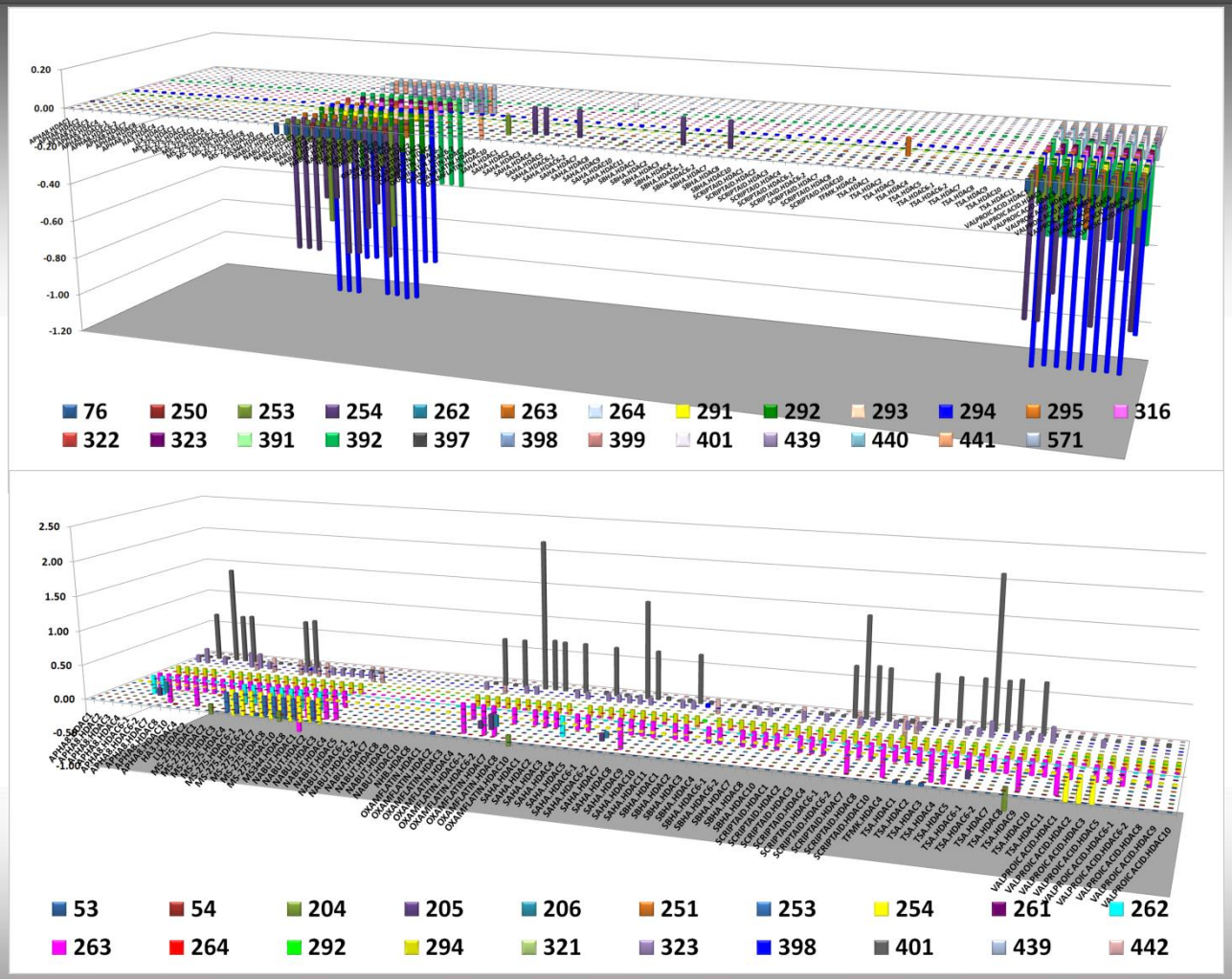






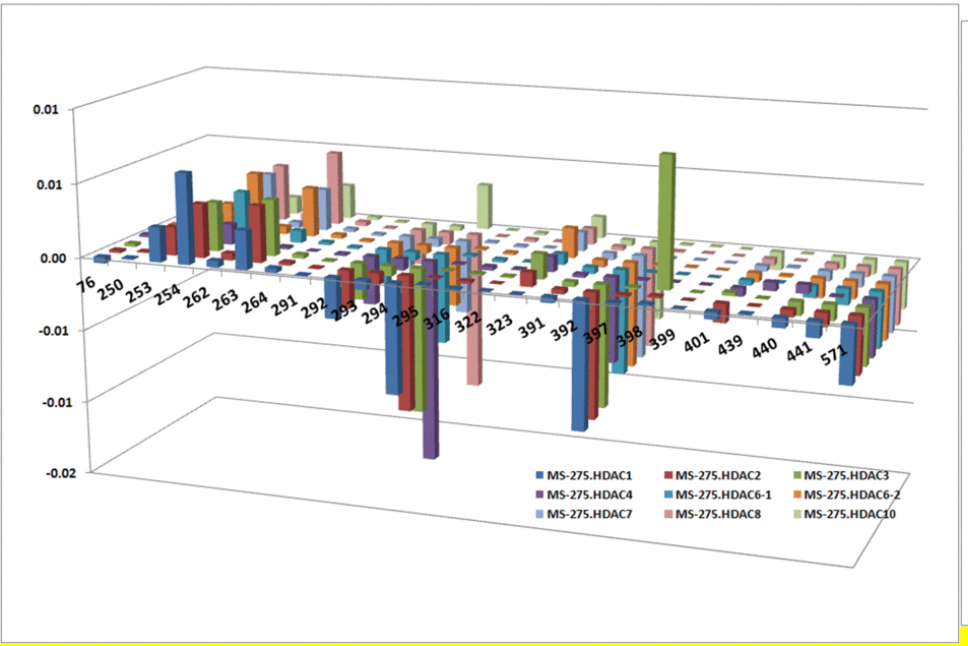
TSA.HDAC7

COMBINEr Model Interpretation

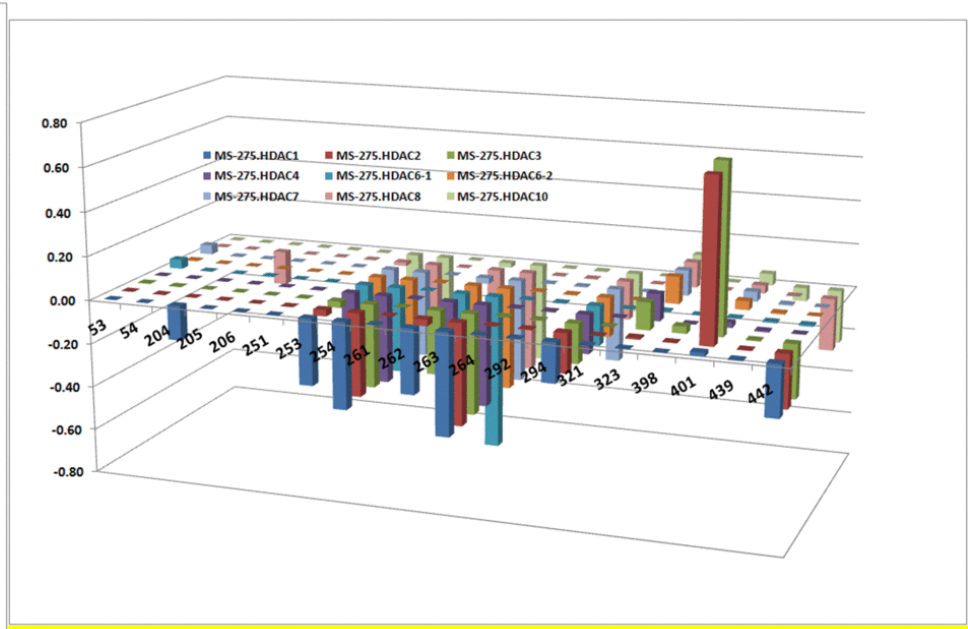


SB 3-D QSAR

Activity Contribution Istograms

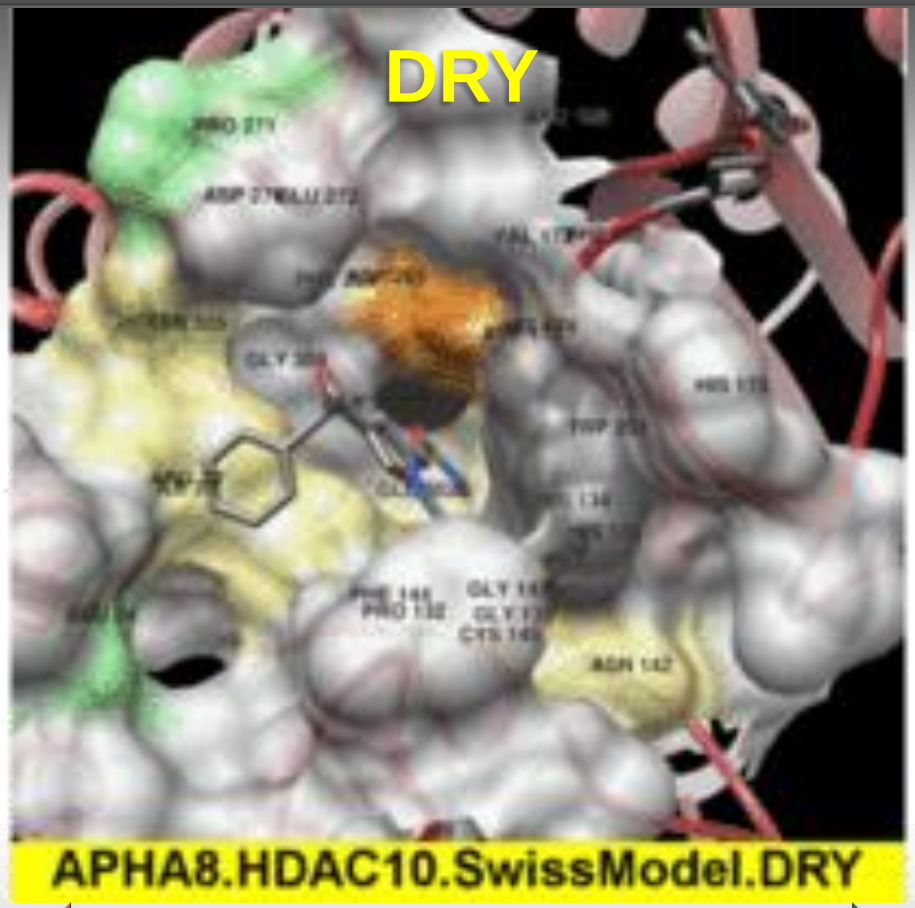


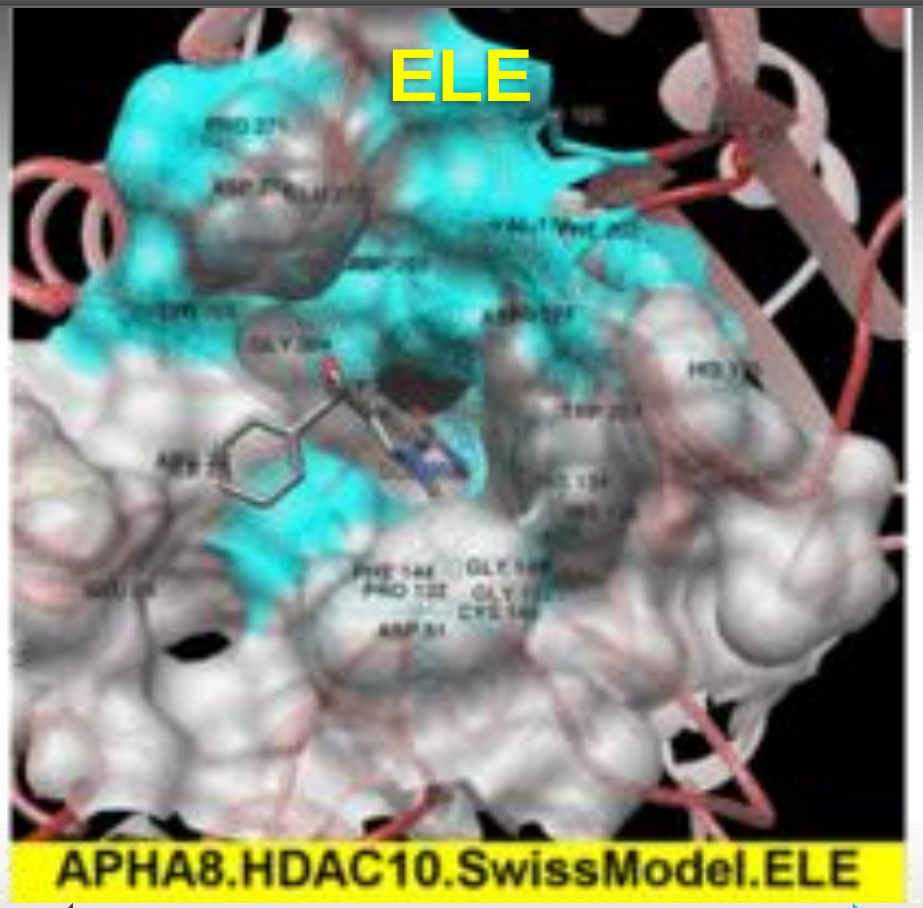
MS275.ELE



MS275.DRY

COMBINEr Model Interpretation





Decreasing Activity | Increasing Activity

Decreasing Activity | Increasing Activity

Activity Contribution Plots

Proteochemometric

Proteochemometric (PCM) modelling is a computational method to model the bioactivity of multiple ligands against multiple related protein targets simultaneously.

PCM modeling can be conceptualized as an extension of QSAR modeling that exploits chemogenomic data by performing a quantitative evaluation of ligand and target structural similarities. As a result, this technique allows the simultaneous navigation, inter- and extrapolation in both chemical space (i.e. ligands) and biological space (i.e. protein target). By the explicit combination of target and ligand information in a single model PCM is capable to analyze and predict SmARs (Structure-multiple Activity Relationships) of a set of compounds.

Proteochemometric

MedChemComm

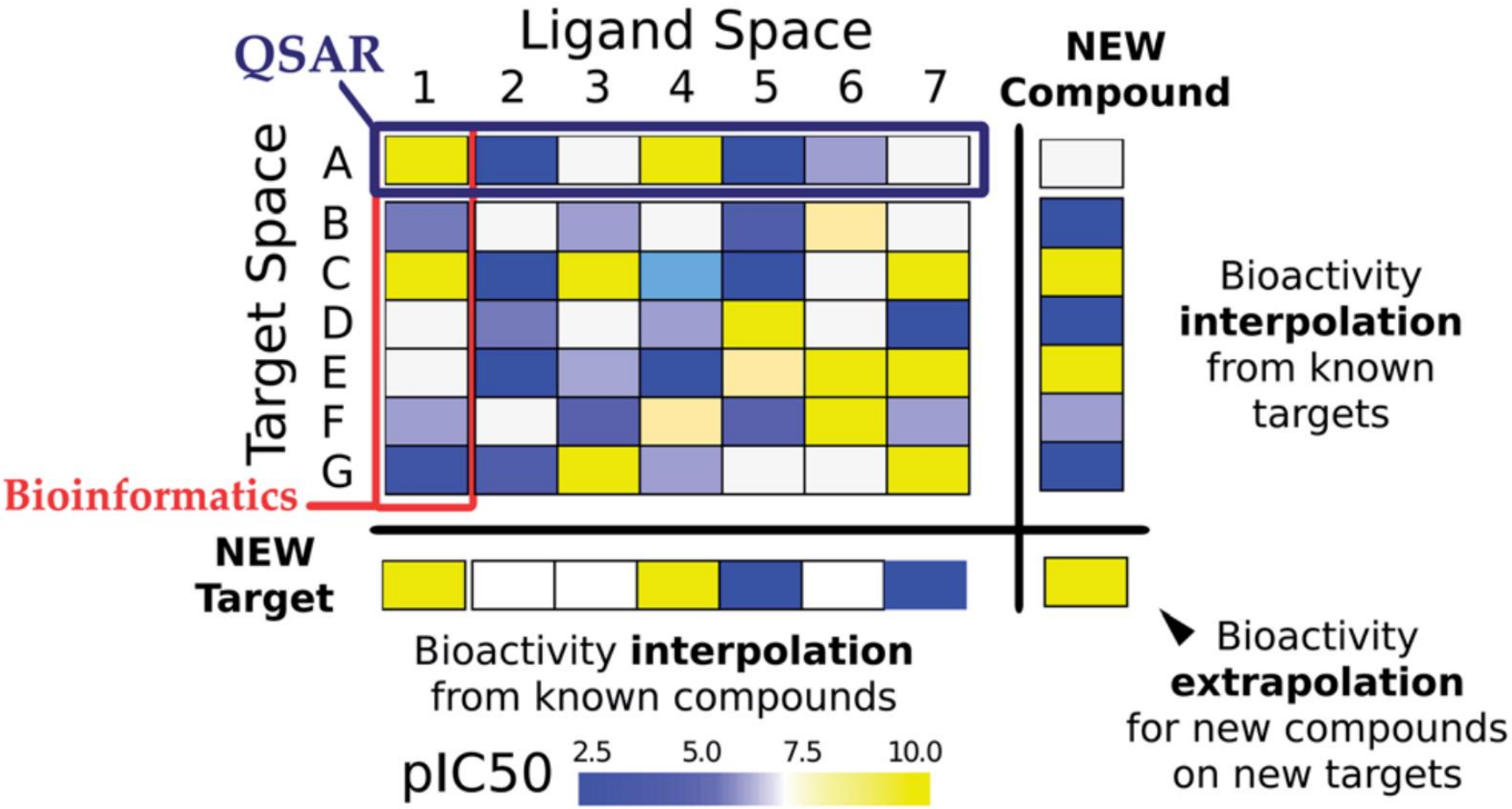
REVIEW

View Article Online
View Journal | View Issue

Cite this: *Med. Chem. Commun.*, 2015, 6, 24

Polypharmacology modelling using proteochemometrics (PCM): recent methodological developments, applications to target families, and future prospects

Proteochemometric



Proteochemometric

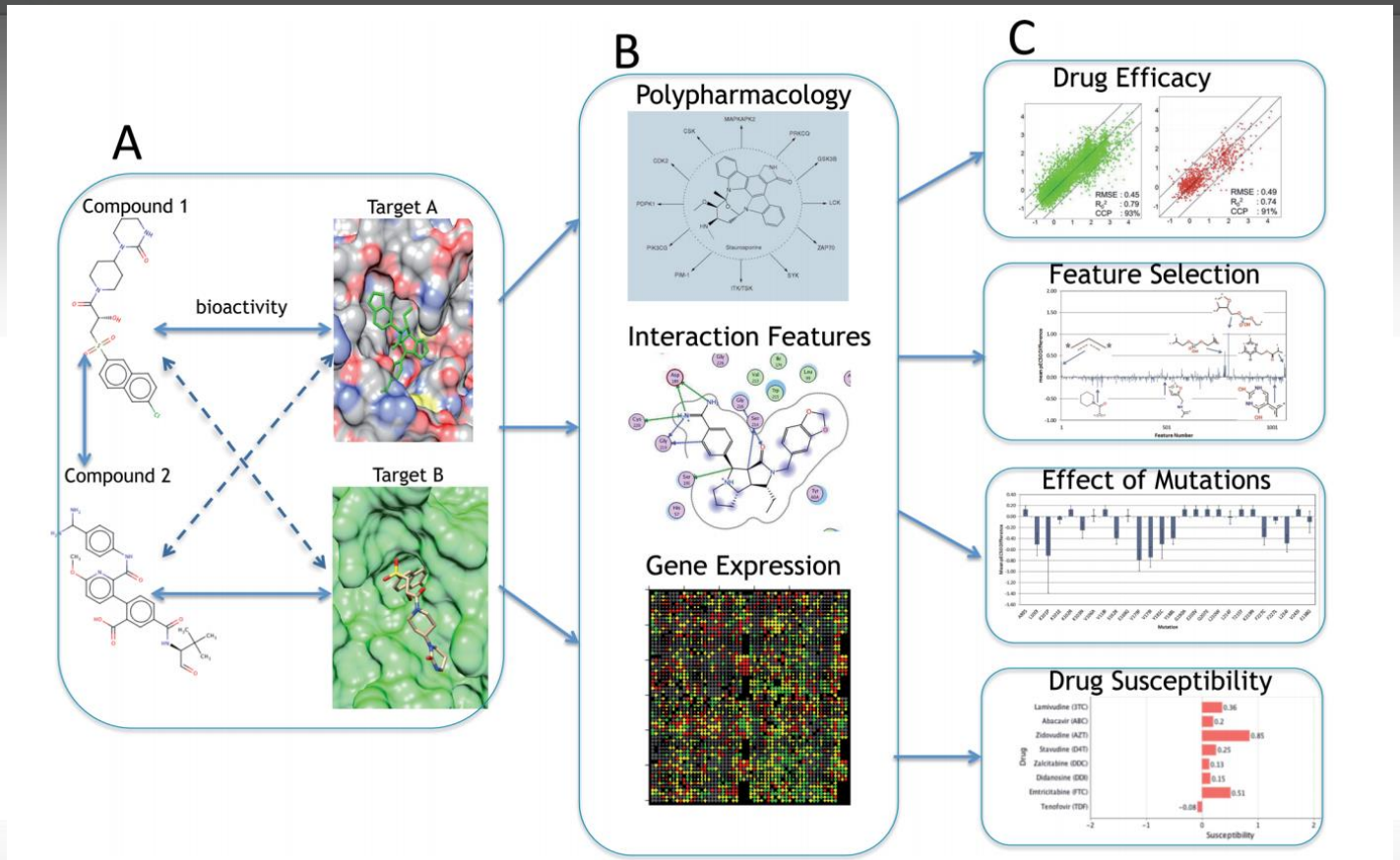


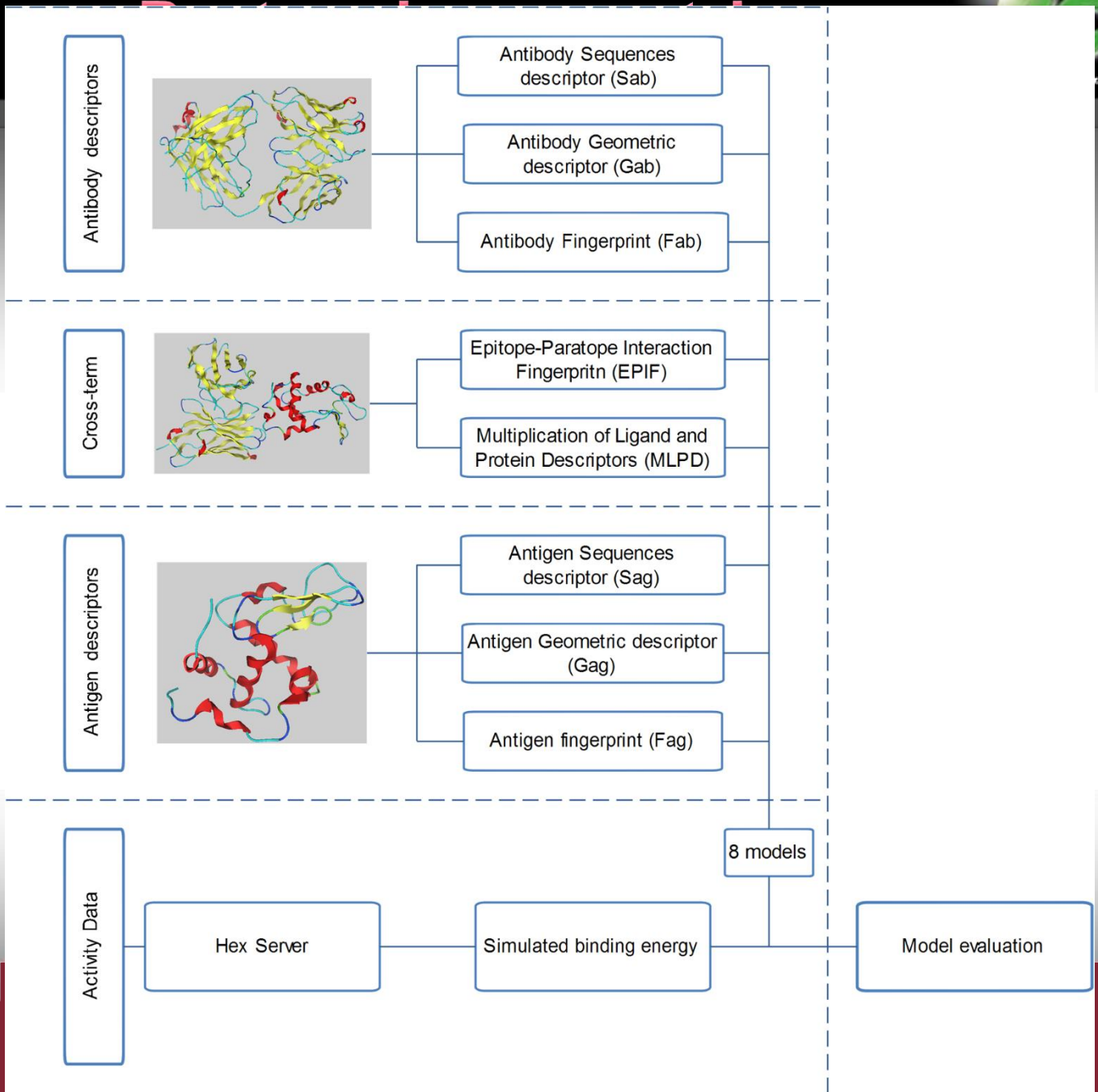
Fig. 2 A systematic overview of proteochemometric modelling. (A) shows the similarity between ligands and drug targets and the utilization of both types of information in PCM. (B) is the representation of different types of input features of ligand and target space (shared bioactivity profiles of ligands, binding pocket residues, gene expression in cell lines, mutational stability, *etc.*) which could be employed in a PCM model depending on the type of output variable. The third block (C) shows the various possible applications of PCM models including measurement of drug efficacy and susceptibility, effect of mutations on activity and compound–target feature selection.

Proteochemometric

Proteochemometric Modeling of the Antigen-Antibody Interaction: New Fingerprints for Antigen, Antibody and Epitope-Paratope Interaction

Tianyi Qiu, Han Xiao, Qingchen Zhang, Jingxuan Qiu, Yiyan Yang, Dingfeng Wu, Zhiwei Cao, Ruixin Zhu

Published: April 22, 2015 • https://doi.org/10.1371/journal.pone.0122416


Antibody descriptors
Antibody Sequences descriptor (Sab)
Antibody Geometric descriptor (Gab)
Antibody Fingerprint (Fab)
Cross-term
Epitope-Paratope Interaction Fingerpritn (EPIF)
Multiplication of Ligand and Protein Descriptors (MLPD)
Antigen descriptors
Antigen Sequences descriptor (Sag)
Antigen Geometric descriptor (Gag)
Antigen fingerprint (Fag)
8 models
Activity Data
Hex Server
Simulated binding energy
Model evaluation

What to do if 3-D Structure of the Target is not known?

Homology Models

Homology Modeling