SAPIENZA
UNIVERSITÀ DI ROMA

FACOLTÀ DI FARMACIA E MEDICINA

Flavio Ballante

Doctoral candidate
Dissertation defense

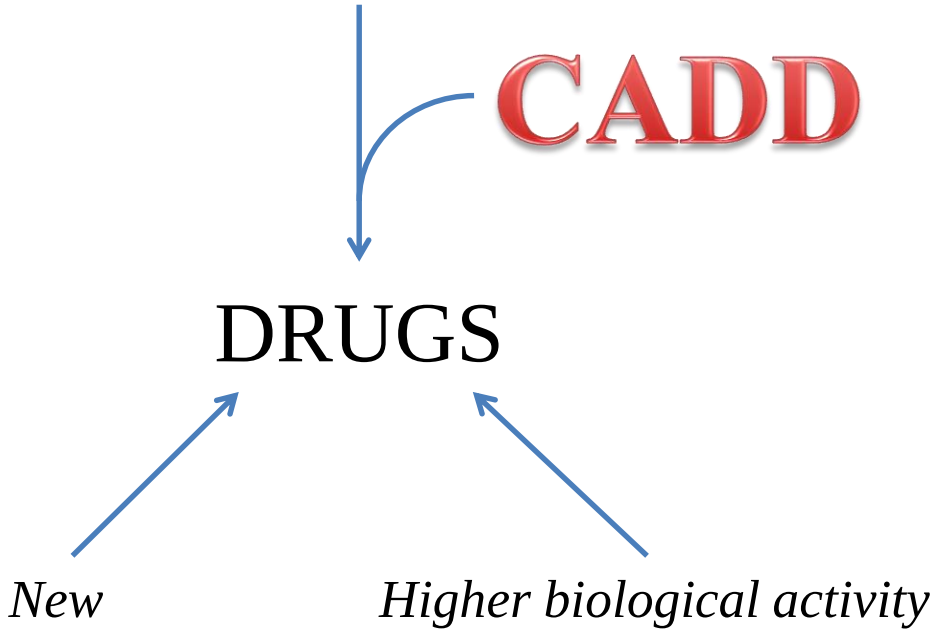
Application of Medicinal Chemistry Methods on Different Classes of Drugs

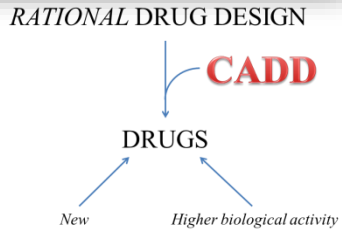
Advisor:
Prof. Rino Ragno

External Examiners:
Prof. Gilbert Kirsch
Prof. Wolfgang Sippl

Dissertation Commitee:
Prof. Daniela Secci
Prof. Daniele Passarella
Prof. Leonardo Guidoni
Prof. Paola B. Arimondo

RATIONAL DRUG DESIGN





CADD

Computer-Aided Drug Design

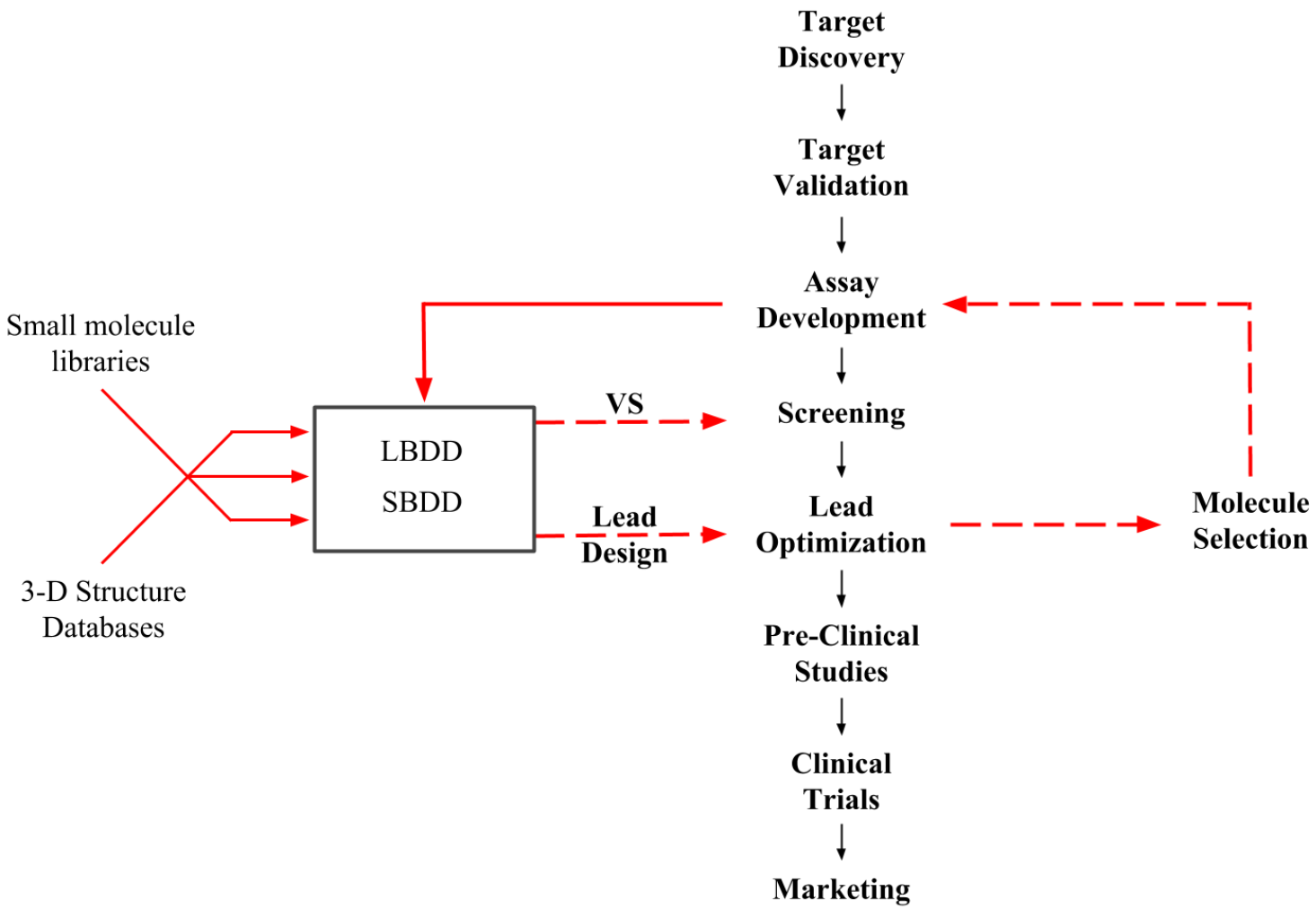
LIGAND BASED (LBDD)

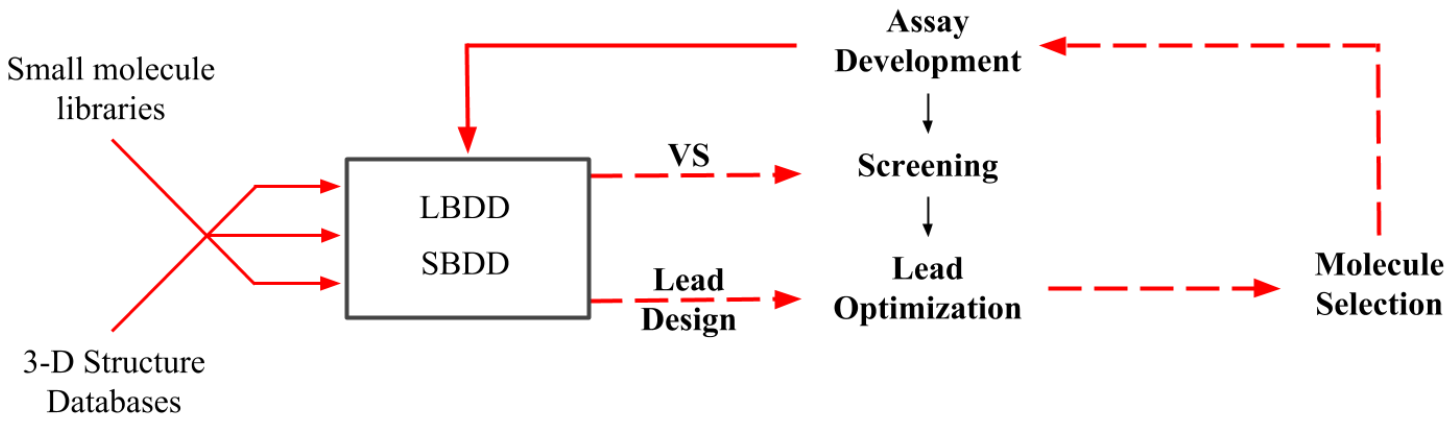
- QSAR
- QSPR
- 3-D QSAR

STRUCTURE BASED (SBDD)

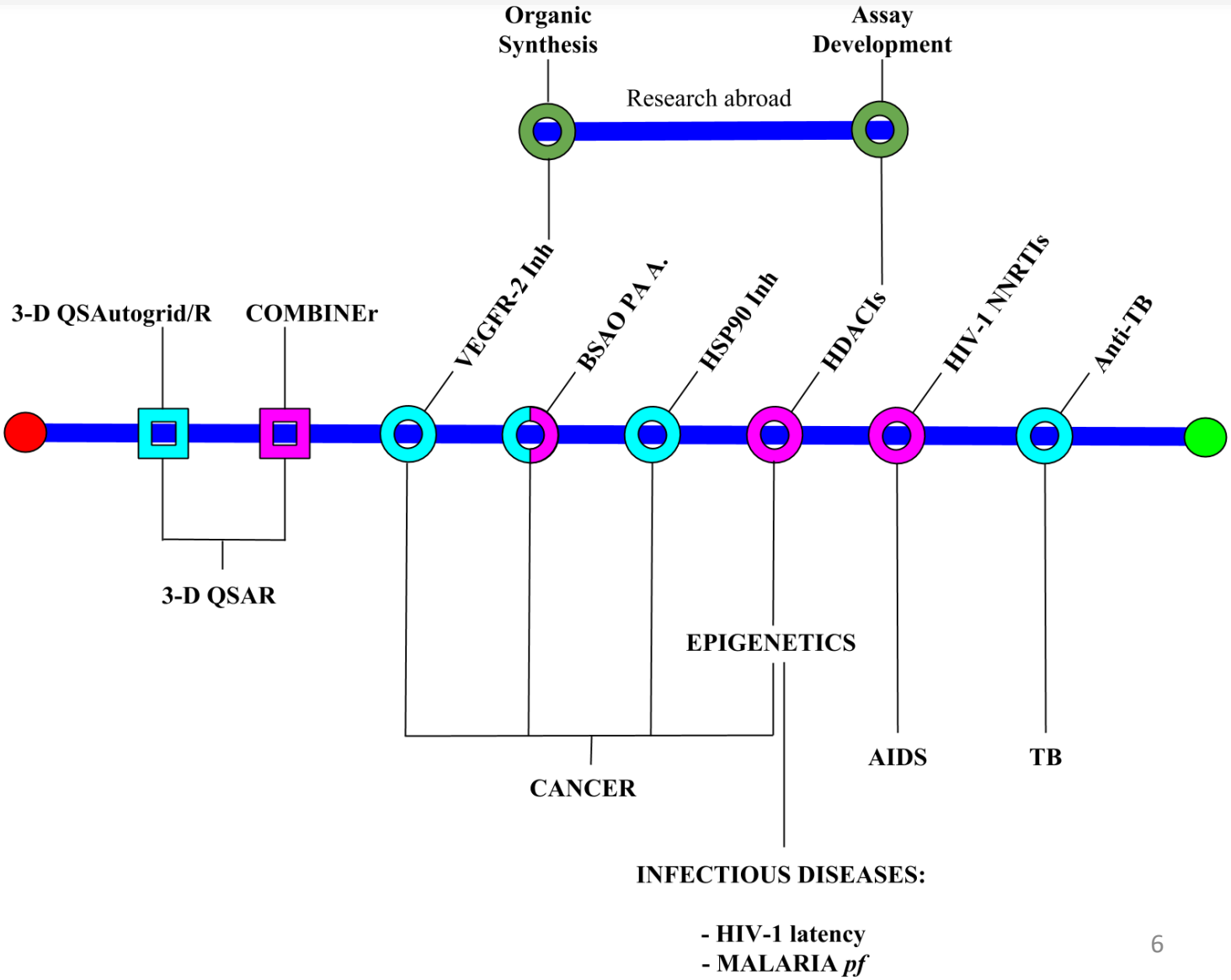
Molecular Docking

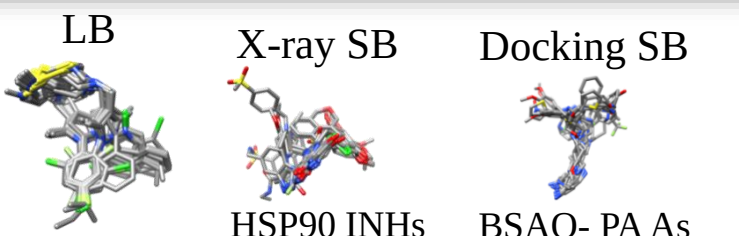
- Protein-Ligand Docking
- Protein-Protein Docking
- Protein-DNA Docking



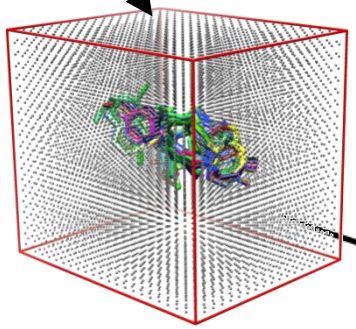


- **Design/validate/optimize different tools to derive quantitative structure-activity relationships (QSAR, 3-D QSAR, and COMBINE) on different molecular classes of current interest.**
- **Highlight/explain the fundamental ligand-protein interactions**
- **Predict untested/novel compounds**
- **Synthetize and evaluate new designed compounds**
- **Optimize existing quantitative models**
- **Obtain new lead compounds**





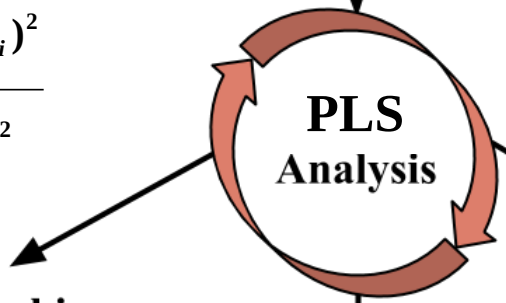
$$SDEP = \sqrt{\frac{\sum_{i=1}^{N}(Y_{exp,i} - Y_{pred,i})^2}{N}}$$



Biological Response

$$r^2 = 1 - \frac{\sum_{i=1}^{N}(Y_{exp,i} - Y_{calc,i})^2}{\sum_{i=1}^{N}(Y_{exp,i} - \bar{Y})^2}$$

$$q^2 = 1 - \frac{\sum_{i=1}^{N}(Y_{exp,i} - Y_{pred,i})^2}{\sum_{i=1}^{N}(Y_{exp,i} - \bar{Y})^2}$$



Graphic Results

Numerical Results

Prediction

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

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*

Quant Struct-Act Rel 1993, 12, (1), 9-20

Generating Optimal Linear Pls Estimations (Golpe) - an Advanced Chemometric Tool for Handling 3d-Qsar Problems.

Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi,S.

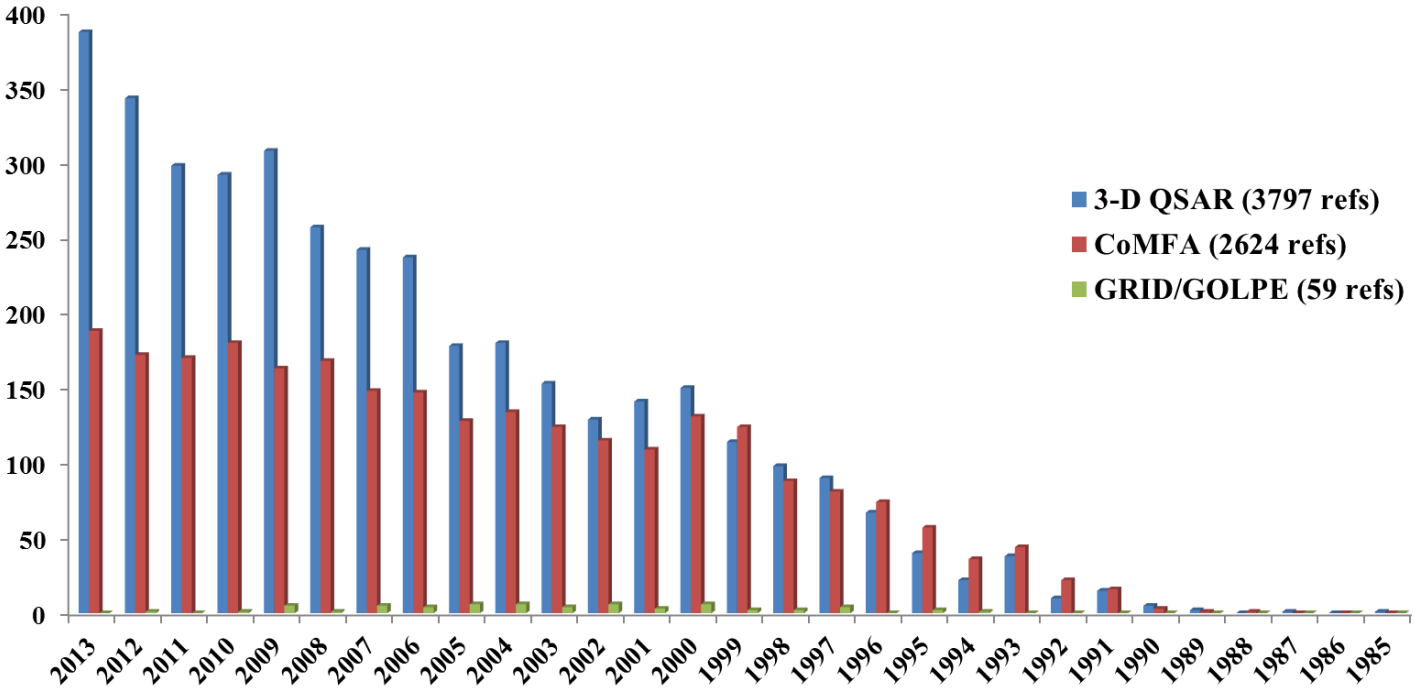
J. Med. Chem. **1994**, *37*, 2589–2601

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

- **Patent restrictions until June 17th 2011**
- **Not automatized proprietary softwares**
- **Still the same procedures**



JOURNAL OF CHEMICAL INFORMATION AND MODELING
Article
pubs.acs.org/jcim

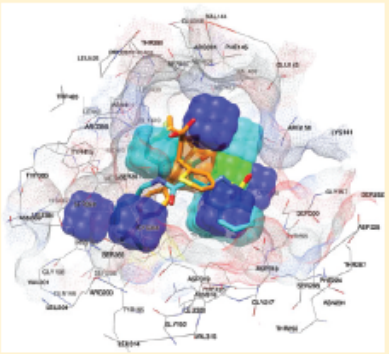
3-D QSAutogrid/R: An Alternative Procedure To Build 3-D QSAR Models. Methodologies and Applications

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

[†]Rome Center for Molecular Design, Dipartimento di Chimica e Tecnologie del Farmaco, Sapienza Università di Roma, P. le A. Moro 5, 00185, Rome, Italy

Supporting Information

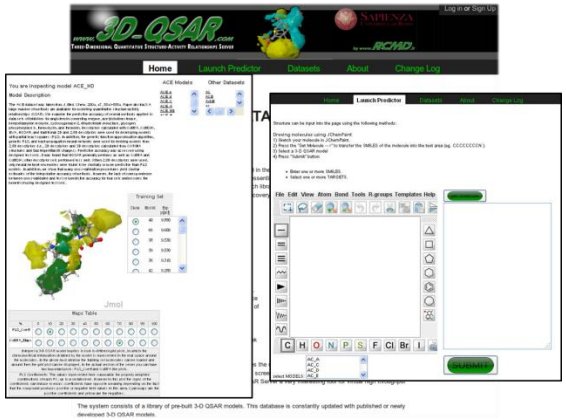
ABSTRACT: Since it first appeared in 1988 3-D QSAR has proved its potential in the field of drug design and activity prediction. Although thousands of citations now exist in 3-D QSAR, its development was rather slow with the majority of new 3-D QSAR applications just extensions of CoMFA. An alternative way to build 3-D QSAR models, based on an evolution of software, has been named 3-D QSAutogrid/R and has been developed to use only software freely available to academics. 3-D QSAutogrid/R covers all the main features of CoMFA and GRID/GOLPE with implementation by multiprobe/multiregion variable selection (MPGRS) that improves the simplification of interpretation of the 3-D QSAR map. The methodology is based on the integration of the molecular interaction fields as calculated by AutoGrid and the R statistical environment that can be easily coupled with many free graphical molecular interfaces such as UCSF-Chimera, AutoDock Tools, JMol, and others. The description of each R package is reported in detail, and, to assess its validity, 3-D QSAutogrid/R has been applied to three molecular data sets of which either CoMFA or GRID/GOLPE models were reported in order to compare the results. 3-D QSAutogrid/R has been used as the core engine to prepare more that 240 3-D QSAR models forming the very first 3-D QSAR server (www.3d-qsar.com) with its code freely available through R-Cran distribution.



3-D QSAutogrid/R

CoMFA and
GRID/GOLPE features

AutoGrid
CAPP
MPGRS
AC Plots on TestSet compds



Core engine of the first
3-D QSAR server
www.3d-qsar.com

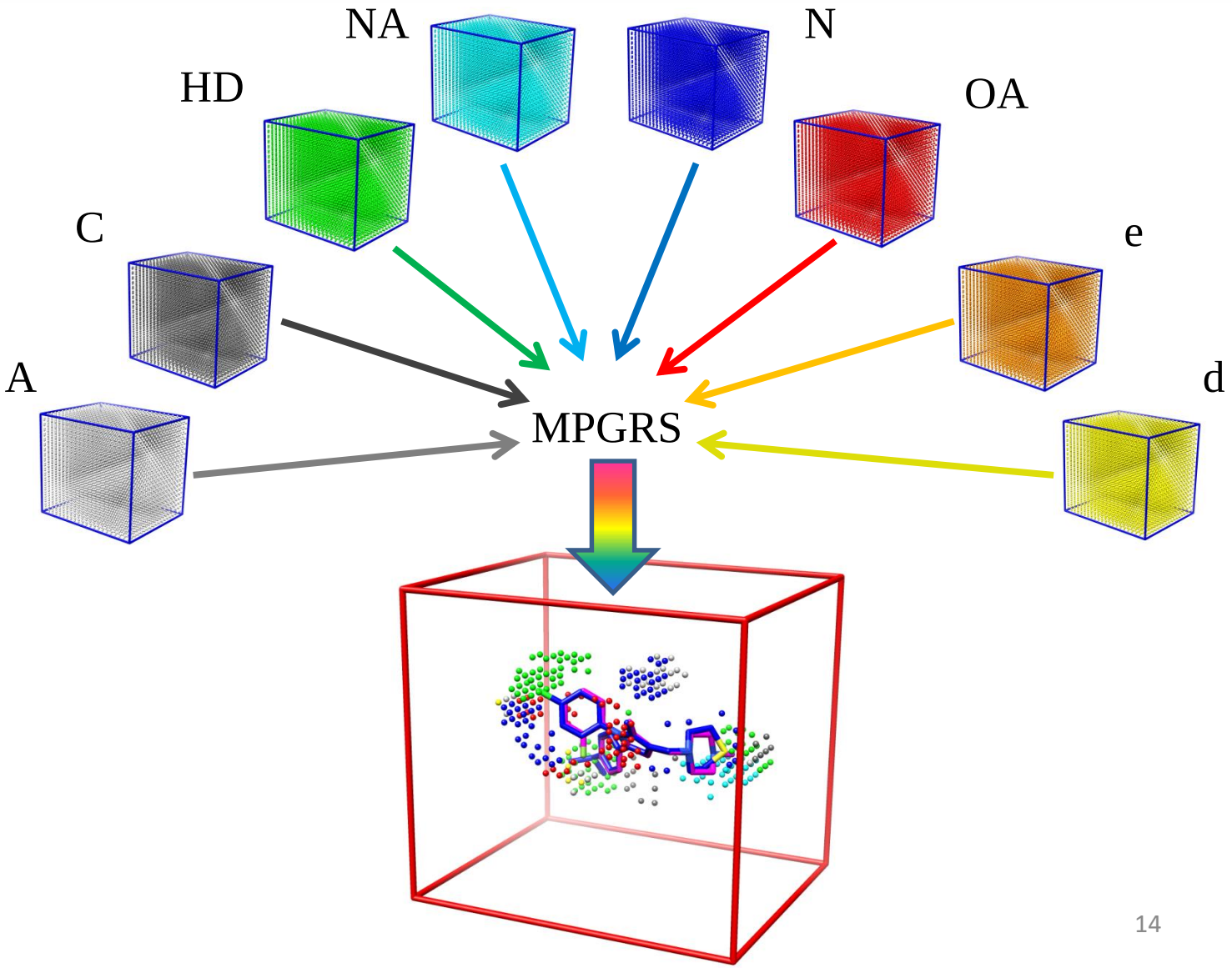
Probe	Description	MPGRS Colour
A	Aromatic Carbon	Gray
C	Aliphatic (sp^3) Carbon	Dark Gray
HD	Hydrogen bonded to heteroatom	Green
NA	Hydrogen-bond-accepting amine nitrogen	Cyan
N	Amide nitrogen	Blue
OA	Hydrogen-bond-accepting oxygen	Red
e	Electrostatic	Orange
d	Desolvation	Yellow

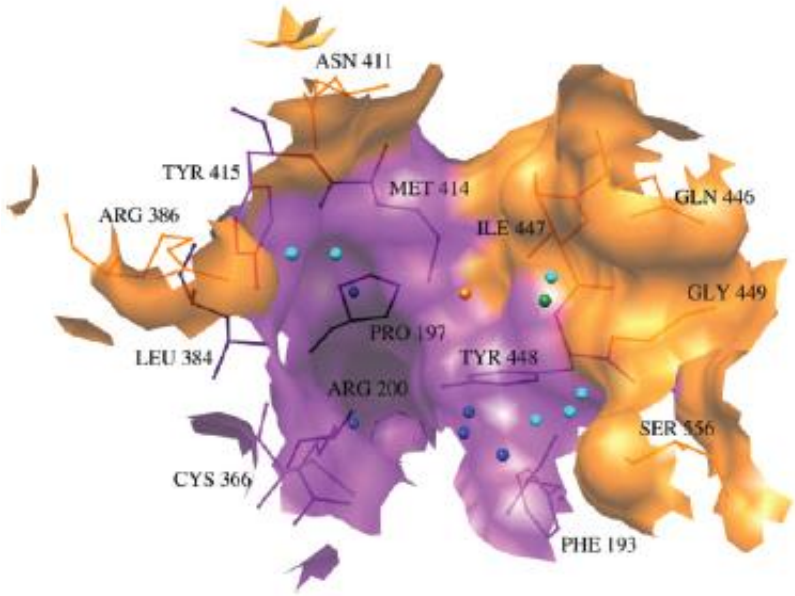
List of the AutoGrid probes employed for MIF calculation and MPGRS subregion color coding.

$$r^2 = 1 - \frac{\sum_{i=1}^{N}(Y_{\exp,i} - Y_{calc,i})^2}{\sum_{i=1}^{N}(Y_{\exp,i} - \overline{Y})^2}$$

$$q^2 = 1 - \frac{\sum_{i=1}^{N}(Y_{\exp,i} - Y_{pred,i})^2}{\sum_{i=1}^{N}(Y_{\exp,i} - \overline{Y})^2}$$

$$SDEP = \sqrt{\frac{\sum_{i=1}^{N}(Y_{\exp,i} - Y_{pred,i})^2}{N}}$$





➢ MPGRS allowed focusing on the most informative regions around the ligands and used all the probes together to reduce the chances of missing important correlations when using single probe 3-D QSARs.

➢ 3-D QSAutogrid/R validation:

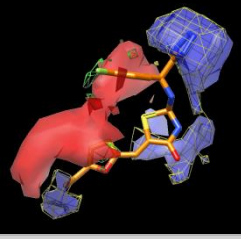
- a data set of aligned opioid-receptor antagonists (LB data set)
- two data sets of HCV NS5B allosteric inhibitors (SB data sets)

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



➢ 3-D QSAutogrid/R validation:

- a data set of aligned opioid-receptor antagonists (LB data set)

3-D QSAutogrid/R

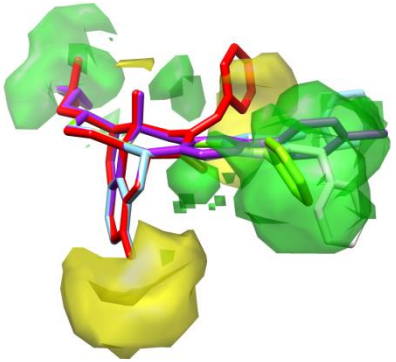
OR	P	PC	r^2	q^2_{LOO}	q^2_{K5FCV}	r^2_{YS}	q^2_{YS}
δ	Autogrid DP	3	0.83	0.70	0.67	0.41	-0.50
μ	Autogrid DP	4	0.85	0.65	0.63	0.52	-0.53
κ	Autogrid DP	3	0.84	0.67	0.63	0.50	-0.53

CoMFA

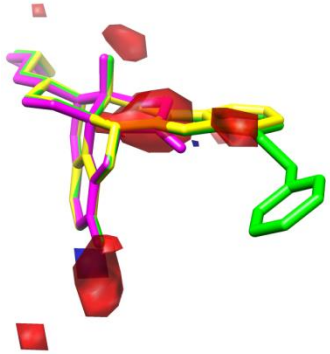
OR	P	PC	r^2	q^2_{LOO}	q^2_{K5FCV}
δ	CoMFA	4	0.91	0.69	-
μ	CoMFA	4	0.92	0.67	-
κ	CoMFA	6	0.96	0.60	-

➢ 3-D QSAutogrid/R validation:

- a data set of aligned opioid-receptor antagonists (LB data set)

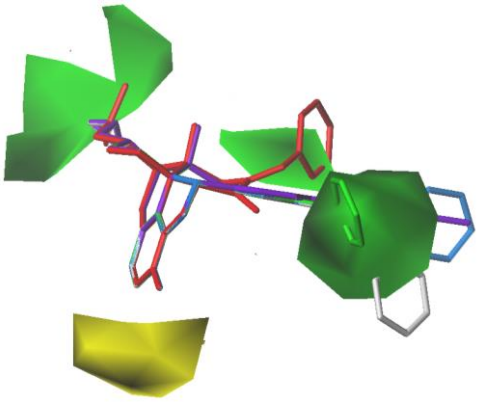


3-D QSAutogrid/R

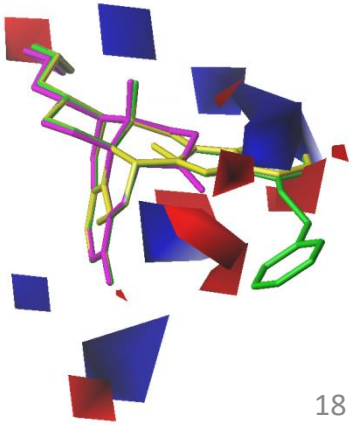


Steric

Electrostatic



CoMFA



➢ 3-D QSAutogrid/R validation:

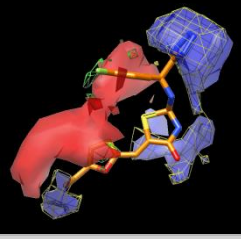
- a data set of aligned opioid-receptor antagonists (LB data set)
- two data sets of HCV NS5B allosteric inhibitors (SB data sets)

662 *J. Chem. Inf. Model.* **2010**, *50*, 662–676

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



➢ 3-D QSAutogrid/R validation:

- two data sets of HCV NS5B allosteric inhibitors (SB data sets)

3-D QSAutogrid/R

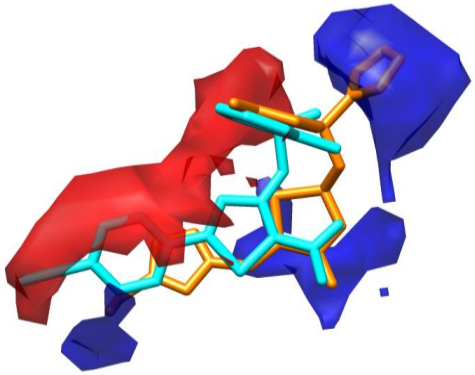
Dataset	P	PC	r^2	q^2_{LOO}	q^2_{K5FCV}	r^2_{YS}	q^2_{YS}
Thumb	Autogrid A	2	0.90	0.67	0.64	0.70	-0.63
Palm	Autogrid A	3	0.96	0.73	0.62	0.68	-1.62

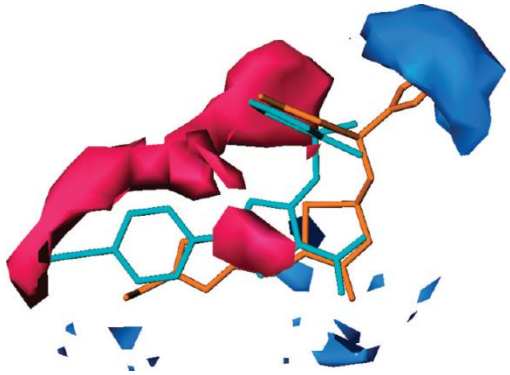
GRID/GOLPE

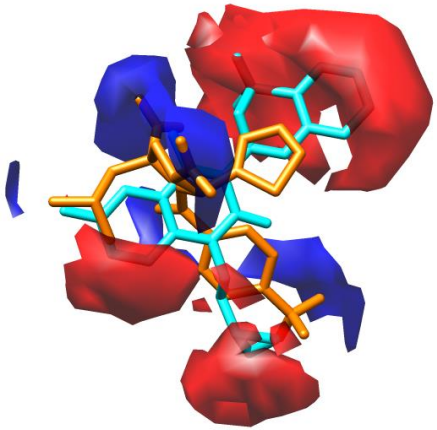
Dataset	P	PC	r^2	q^2_{LOO}	q^2_{K5FCV}	r^2_{YS}	q^2_{YS}
Thumb	GRID/GOLPE/C1=	3	0.99	-	0.69	-	-
Palm	GRID/GOLPE/C1=	3	0.99	-	0.55	-	-

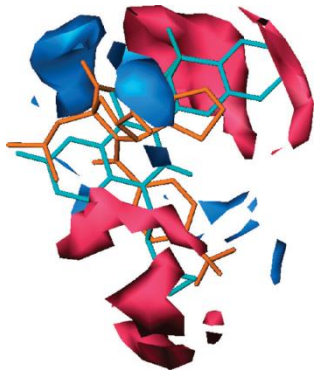
➢ 3-D QSAutogrid/R validation:

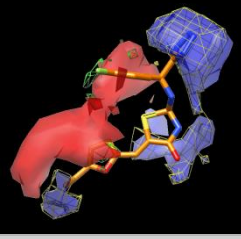
- two data sets of HCV NS5B allosteric inhibitors (SB data sets)











➢ 3-D QSAutogrid/R

- ✓ Successfully validated
- ✓ New automatized 3-D QSAR procedure
- ✓ Completely free for academics
- ✓ New straightforward features : MPGRS and CAPP
- ✓ **Open3DQSAR**[1]

[1] *Journal of Molecular Modeling* **2011**, *17*, 201-208

JOURNAL OF CHEMICAL INFORMATION AND MODELING
Article
pubs.acs.org/jcim

Pharmacophore Assessment Through 3-D QSAR: Evaluation of the Predictive Ability on New Derivatives by the Application on a Series of Antitubercular Agents

Laura Friggeri,[§,†] Flavio Ballante,*[,§,‡] Rino Ragno,*[,‡] Ira Musmuca,[‡] Daniela De Vita,[†] Fabrizio Manetti,[○] Mariangela Biava,[†] Luigi Scipione,[†] Roberto Di Santo,[⊥,†] Roberta Costi,[⊥,†] Marta Feroci,[∥] and Silvano Tortorella[†]

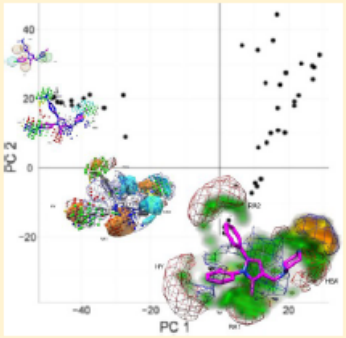
[‡]Rome Center for Molecular Design and [⊥]Istituto Pasteur-Fondazione Cenci Bolognetti, [†]Dipartimento di Chimica e Tecnologie del Farmaco, Sapienza Università di Roma, P. le A. Moro 5, 00185 Rome, Italy

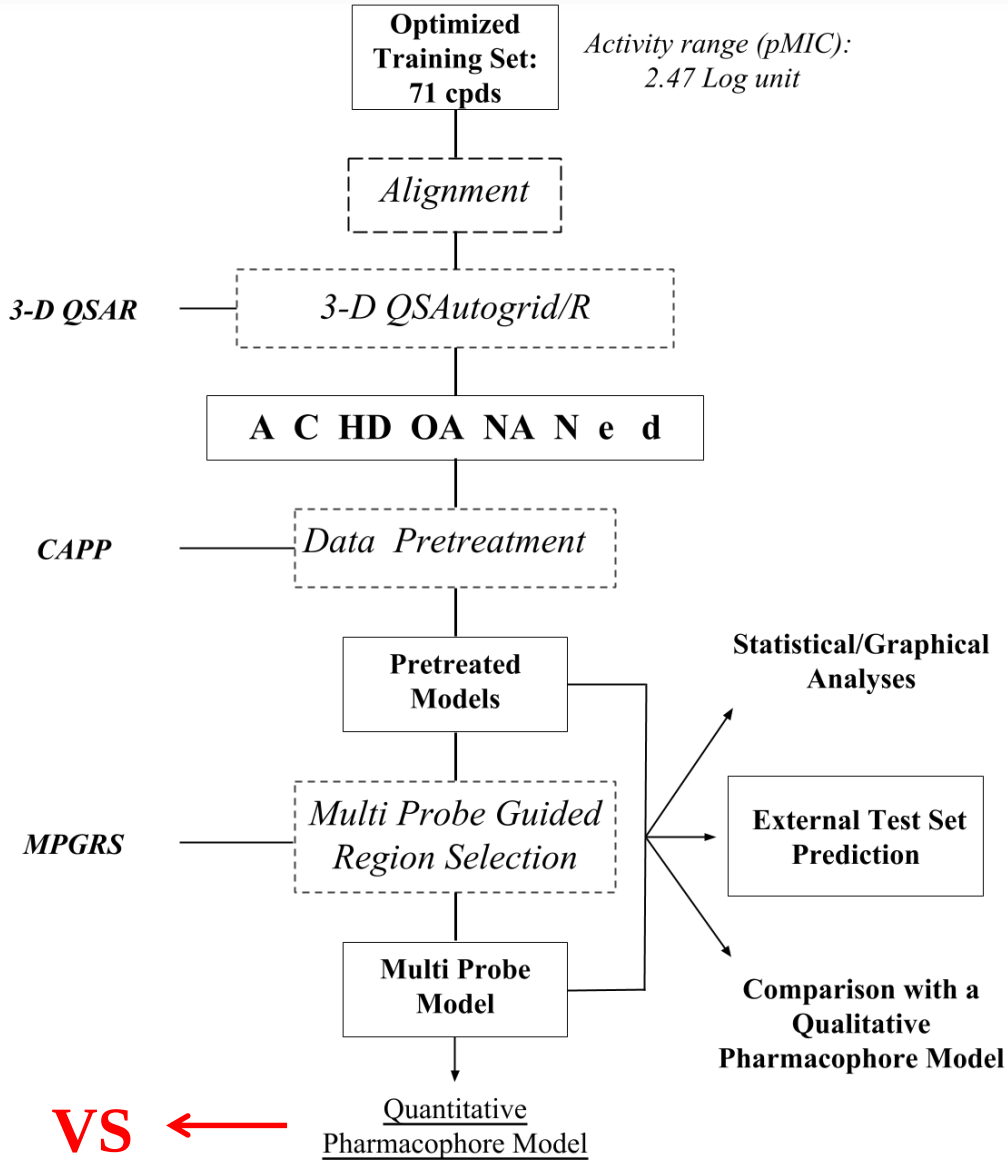
[○]Dipartimento di Biotecnologie, Chimica e Farmacia, Università degli Studi di Siena, Via Aldo Moro 2, I-53100 Siena, Italy

[∥]Dipartimento di Scienze di Base e Applicate per l'Ingegneria, Sapienza Università di Roma, Via Castro Laurenziano 7, I-00161 Rome, Italy

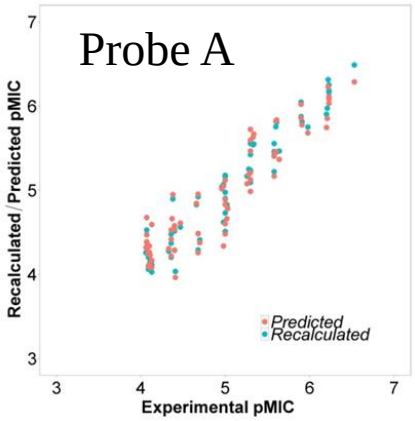
Supporting Information

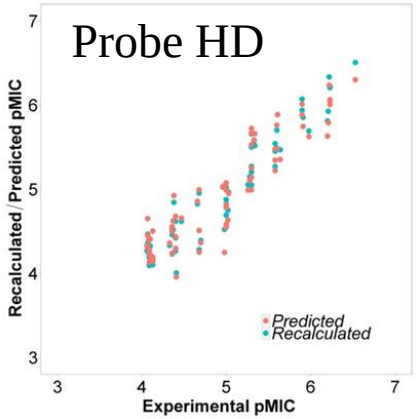
ABSTRACT: Pharmacophoric mapping is a useful procedure to frame, especially when crystallographic receptor structures are unavailable as in ligand-based studies, the hypothetical site of interaction. In this study, 71 pyrrole derivatives active against *M. tuberculosis* were used to derive through a recent new 3-D QSAR protocol, 3-D QSAutogrid/R, several predictive 3-D QSAR models on compounds aligned by a previously reported pharmacophoric application. A final multiprobe (MP) 3-D QSAR model was then obtained configuring itself as a tool to derive pharmacophoric quantitative models. To stress the applicability of the described models, an external test set of unrelated and newly synthesized series of R-4-amino-3-isoxazolidinone derivatives found to be active at micromolar level against *M. tuberculosis* was used, and the predicted bioactivities were in good agreement with the experimental values. The 3-D QSAutogrid/R procedure proved to be able to correlate by a single multi-informative scenario the different activity molecular profiles thus confirming its usefulness in the rational drug design approach.

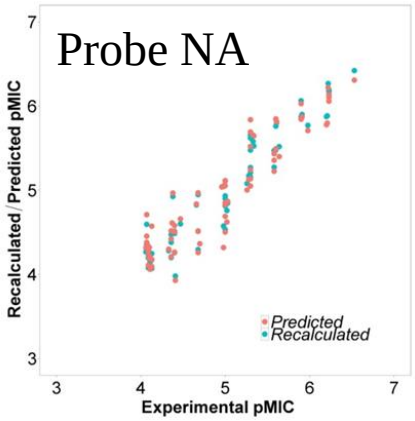


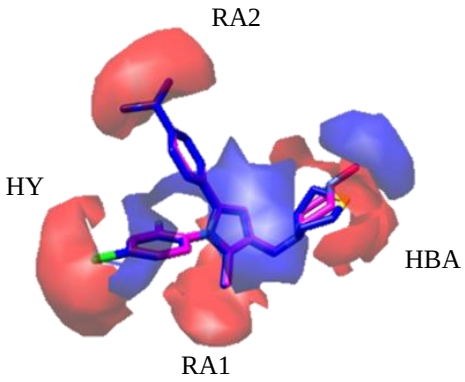


model	P	PC	r^2	q^2_{LOO}	q^2_{K5FCV}	r^2_{YS}	q^2_{YS}	V
1	A	3	0.92	0.86	0.85	0.36	-0.33	3758
2	C	3	0.92	0.86	0.85	0.37	-0.33	4492
3	HD	3	0.91	0.85	0.84	0.39	-0.31	1217
4	NA	3	0.91	0.86	0.85	0.31	-0.33	531
5	N	3	0.91	0.85	0.85	0.32	-0.30	477
6	OA	3	0.91	0.85	0.85	0.36	-0.33	658
7	e	4	0.88	0.78	0.76	0.40	-0.48	468
8	d	4	0.91	0.85	0.84	0.35	-0.44	4412

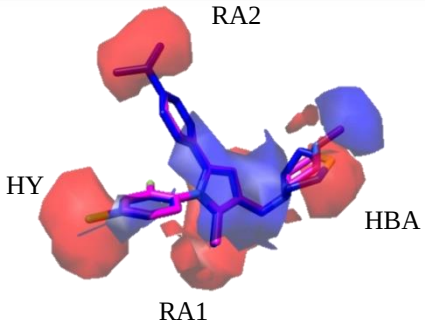




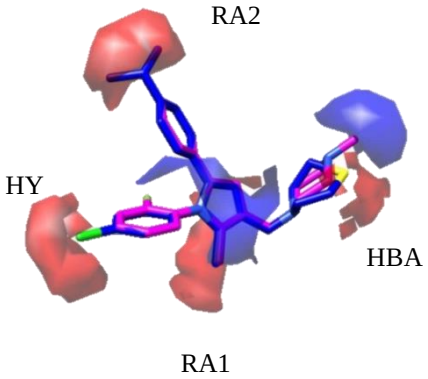




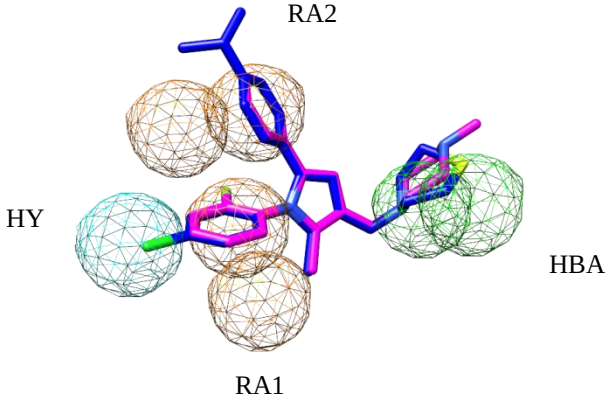
Probe A



Probe HD

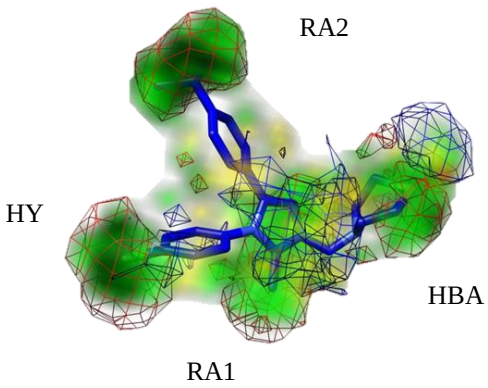


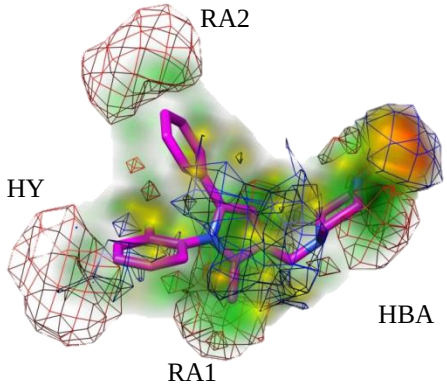
Probe NA



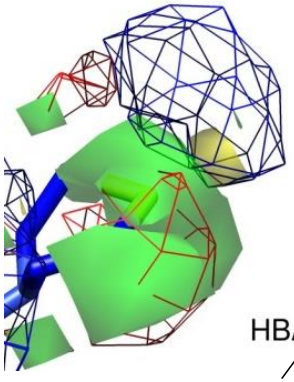
Original pharmacophore obtained by M. Botta group.[1,2]

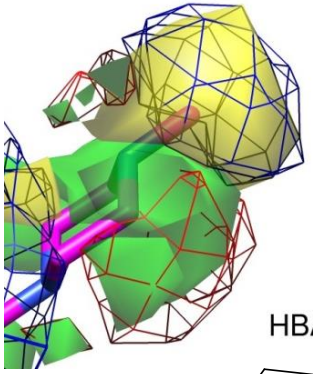
1.Manetti, F.; Corelli, F.; Biava, M.; Fioravanti, R.; Porretta, G. C.; Botta, M., Building a pharmacophore model for a novel class of antitubercular compounds. *Farmaco* **2000**, 55, 484-91.
2. Biava; Mariangela; Fioravanti; Rossella; Porretta; Cesare, G.; Deidda; Delia; Lampis; Giorgio; Pompei; Raffaello; Tafi; Andrea; Manetti; Fabrizio, New derivatives of toluidine: Synthesis, antitubercular activity and pharmacophore hypothesis. *Med. Chem. Res* **2002**.

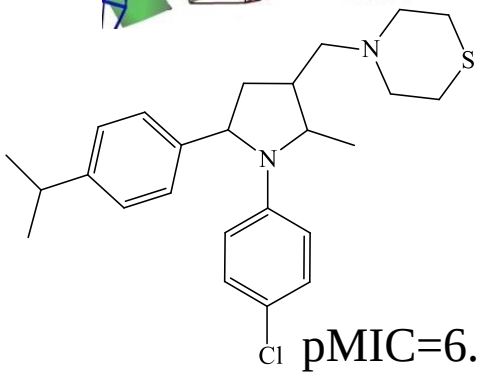




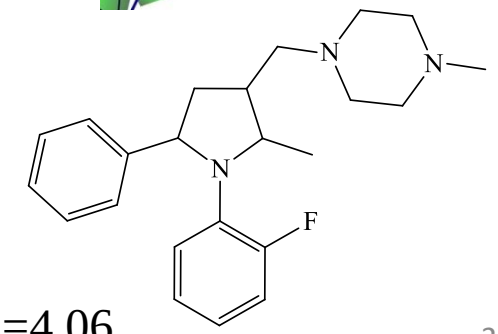
Probe HD



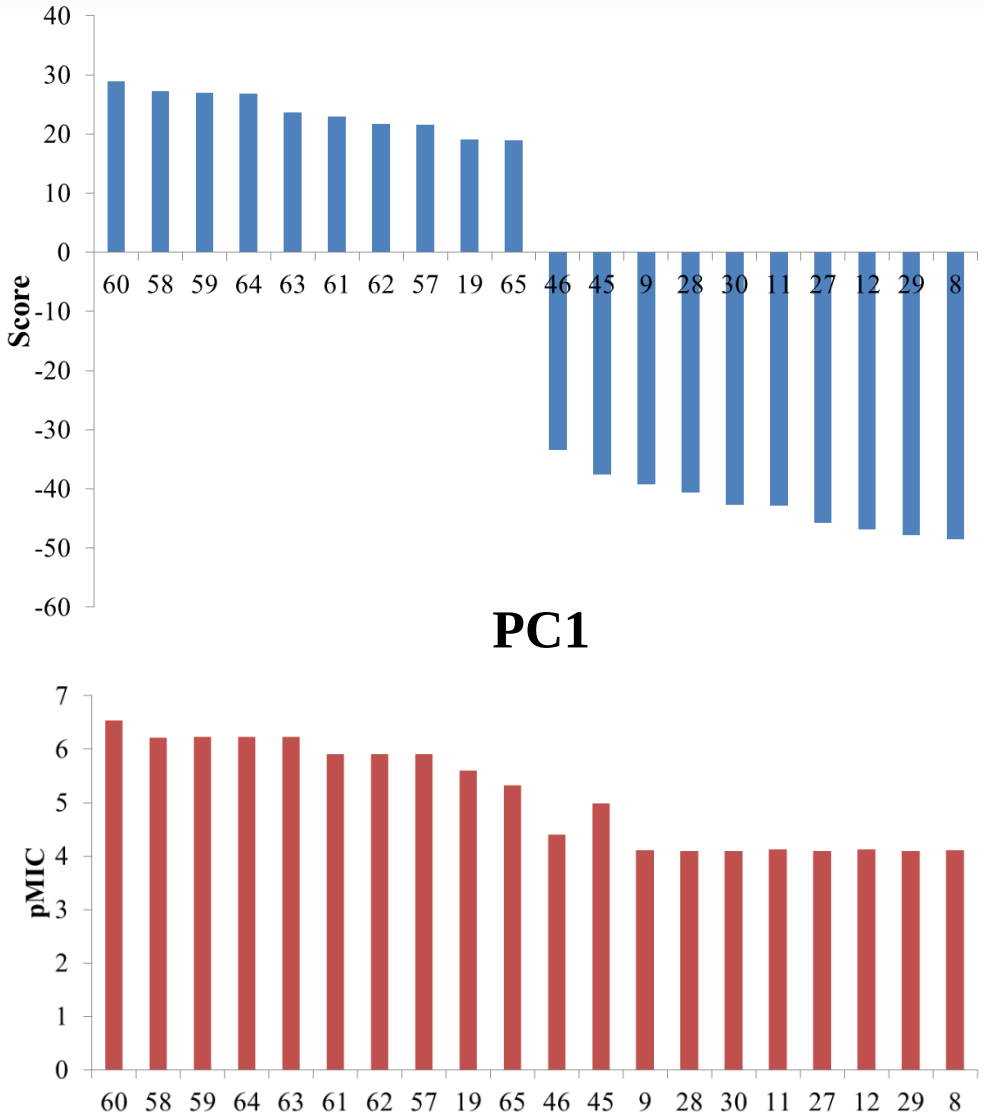


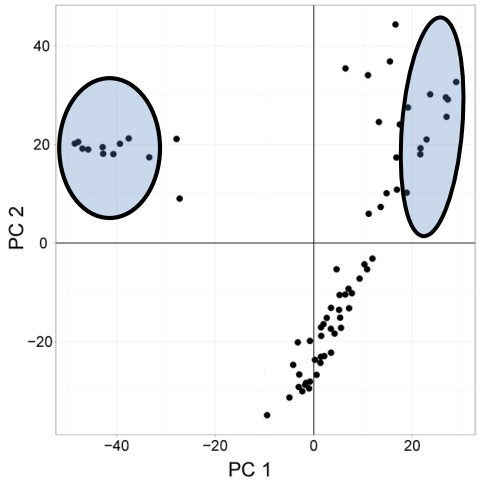


pMIC=6.53

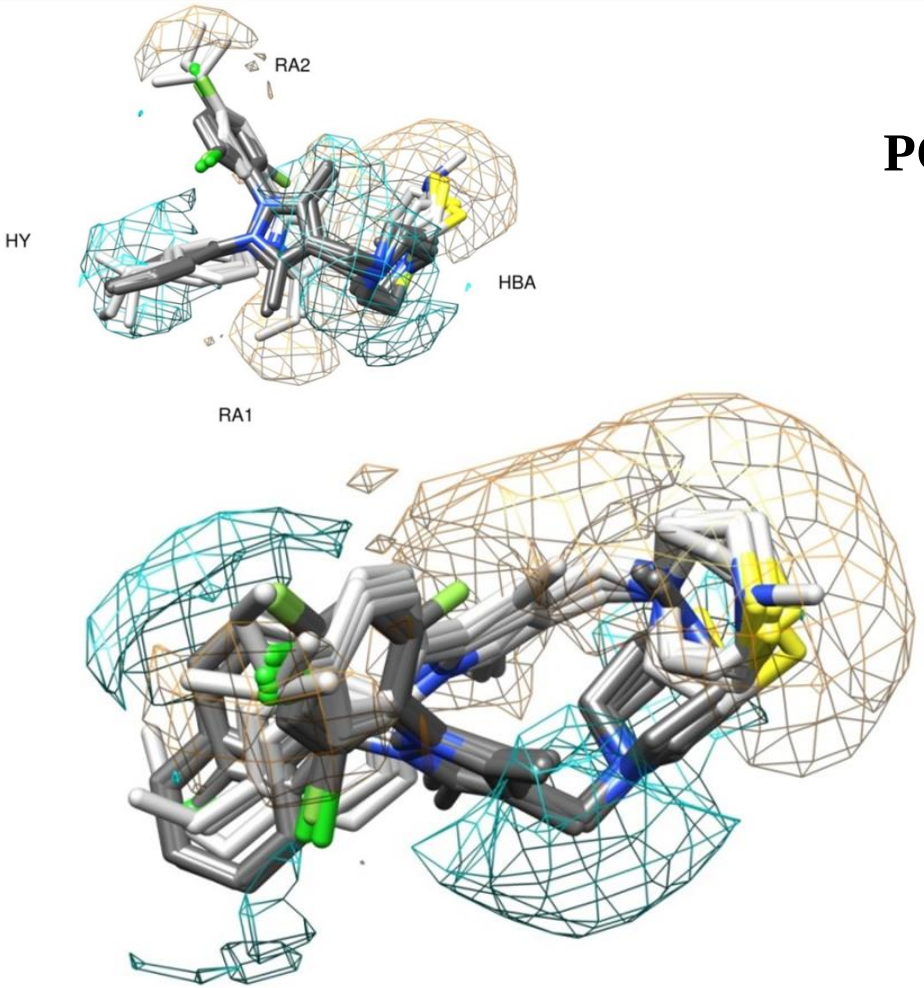


pMIC=4.06



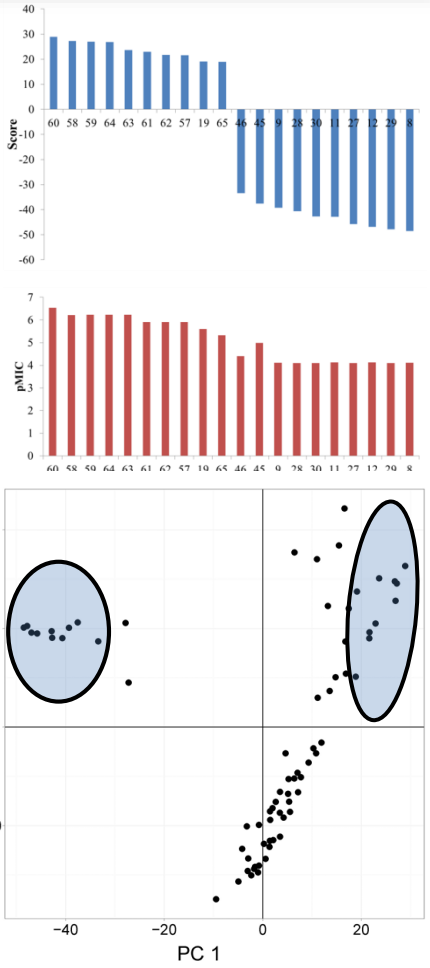


PC1 *vs* PC2 scores plot derived from A probe analysis



PLS-loadings contour maps from the A probe analysis at PC1 (contour levels: 60%; positive: orange, negative: cyan).

PC1

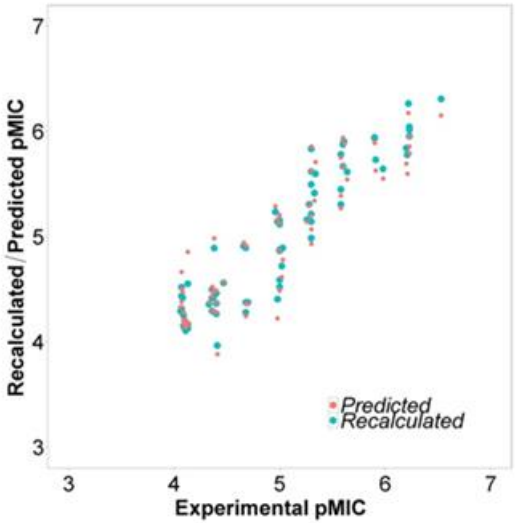


PC1 *vs* PC2 scores plot derived from A probe analysis

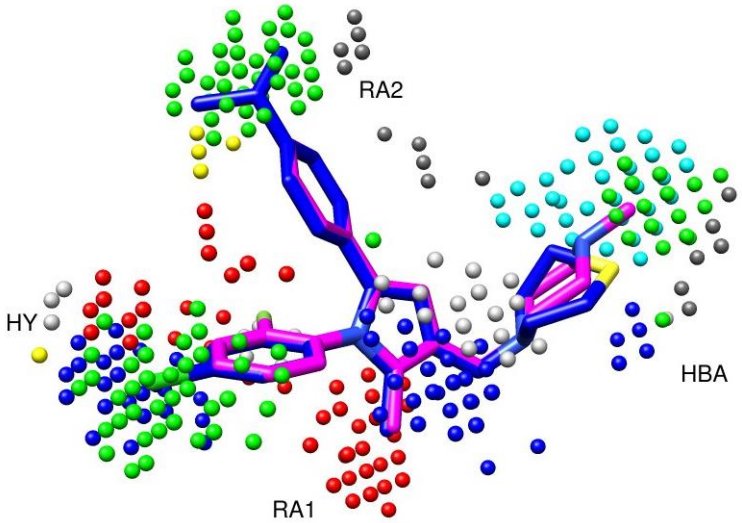
$PC_{FL:SL}$	r^2	q^2_{LOO}	q^2_{K5FCV}	$SDEP_{LOO}$	$SDEP_{K5FCV}$	r^2_{YS}	q^2_{YS}
1:3	0.88	0.80	0.80	0.32	0.32	0.31	-0.31

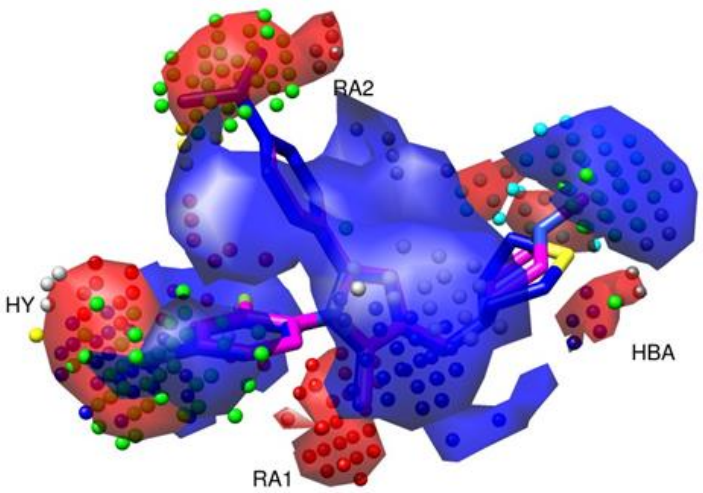
Statistical Results Obtained from MPGRS Analysis.

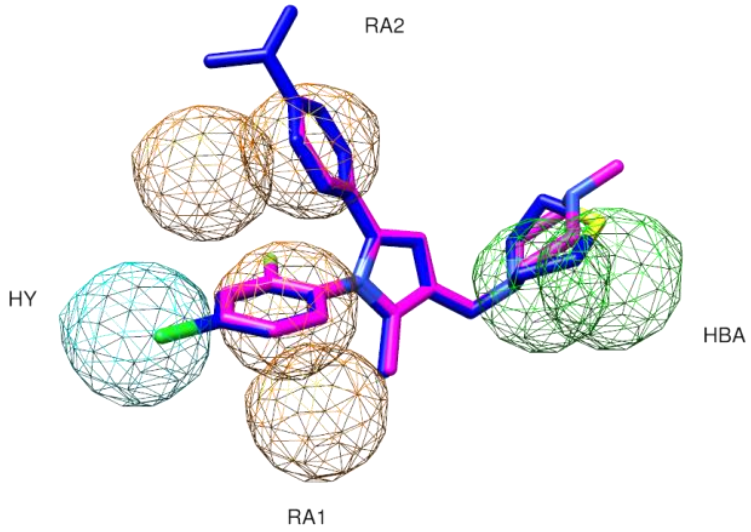
$PC_{FL:SL}$: optimal number of principal first level (FL) and second level (SL) components for the MPGRS model; r^2: conventional square-correlation coefficient; q^2_{LOO}: cross-validation correlation coefficient using the leave-one-out method; q^2_{K5FCV}: cross-validation correlation coefficient using the *k*-fold cross-validation with 5 random groups and 100 iterations; r^2_{YS}: average square correlation coefficient obtained after Y-scrambling process using 100 iterations; q^2_{YS}: average cross-validation correlation coefficient using the leave-one-out method obtained after Y-scrambling process using 100 iterations

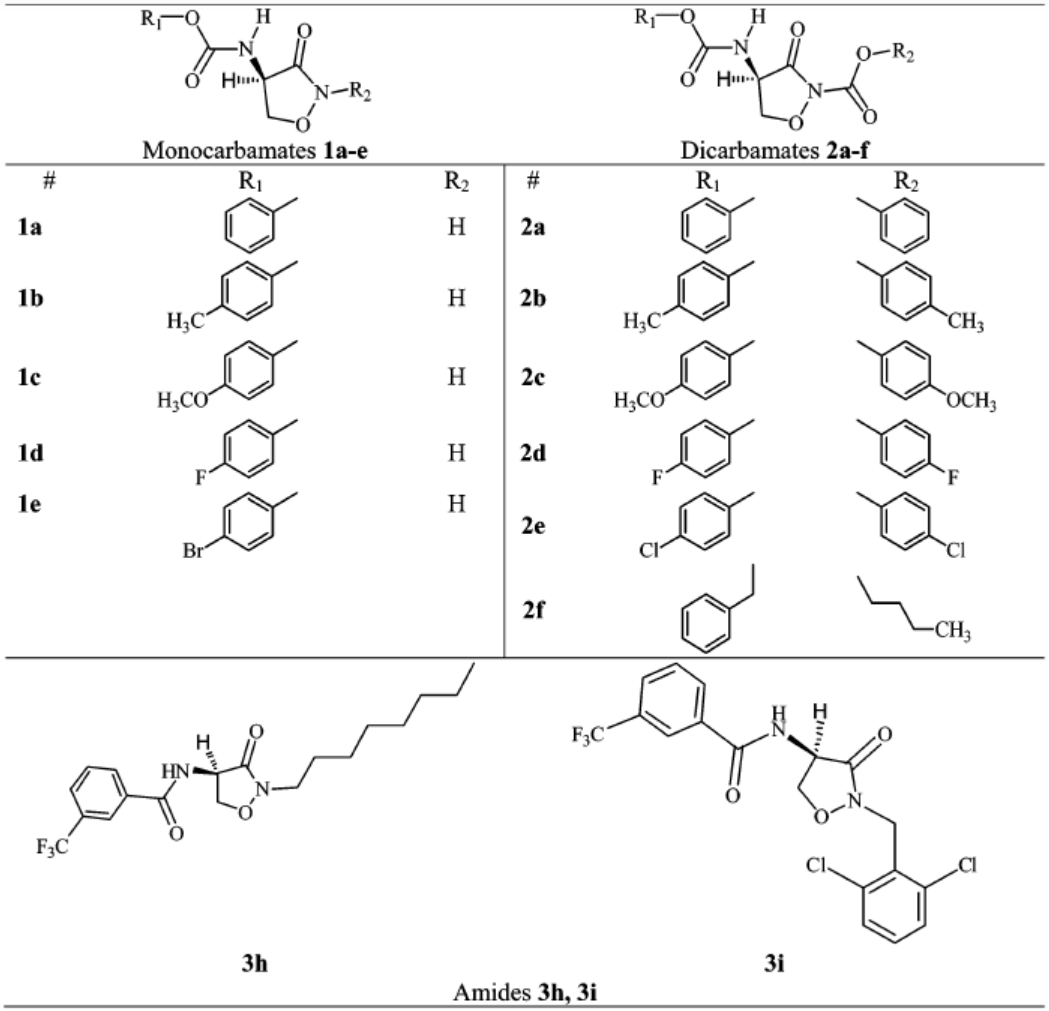


Fitting (r^2) and Cross-Validation (q^2 K-5-Fold) plots at $PC_{1:3}$









External Test Set: R-4-Amino-3-isoxazolidinone Derivatives: Monocarbamates (1a−e), Dicarbamates (2a−f), and Amides (3h,i)

MIC data for D-4-amino-3-isoxazolidinone derivatives

compd	MIC(μg/mL)[a]	pMIC[b]
1a	32	3.84
1b	32	3.87
1c	32	3.90
1d	3.1	4.89
1e	32	3.97
2a	32	4.03
2b	32	4.06
2c	32	4.10
2d	3.1	5.09
2e	32	4.11
2f	32	4.02
3h	64	3.78
3i	32	4.13

[a] *M. tuberculosis* H37Rv (ATCC 27294) was used; MIC values represent the minimal concentrations of compounds completely inhibiting visible growth of mycobacteria.

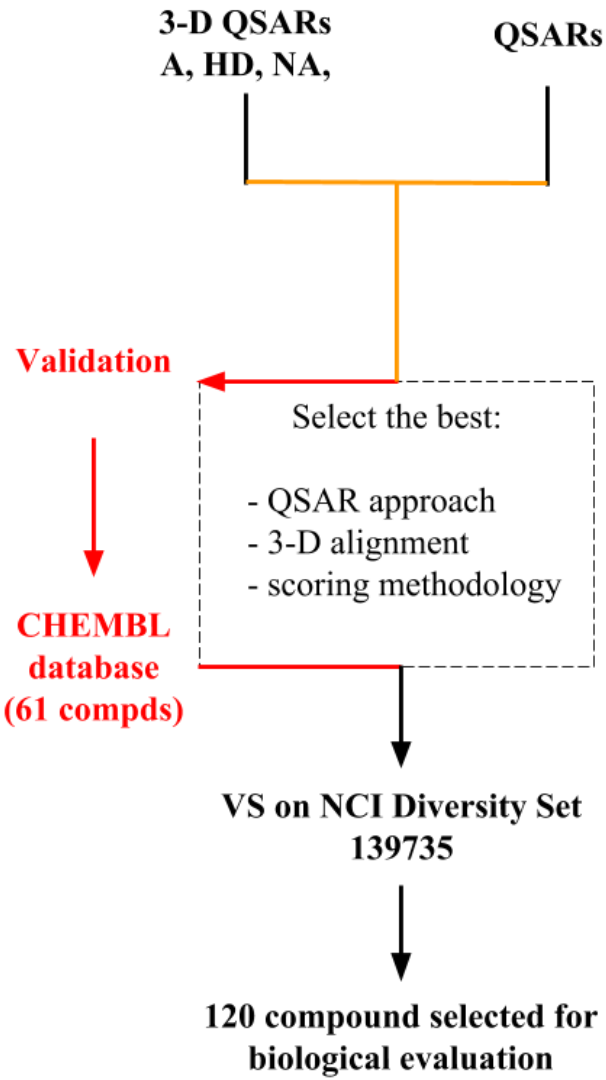
[b] $pMIC = -Log [MIC(\mu M) \times 10^{-6}]$

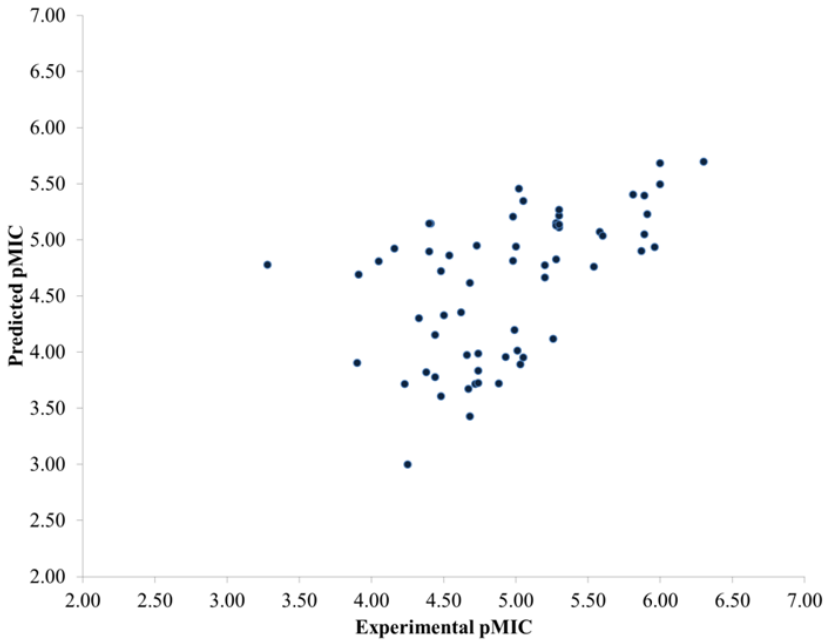
P	PC	$SDEP_{EXT}$
A	3	0.88
C	3	0.88
HD	3	0.81
NA	3	0.82
N	3	0.83
OA	3	0.84
e	4	0.90
d	4	1.51

Test Set predictions: SDEP values considering the optimal PCs; P: AutoGrid Probe; PC: optimal number of principal components/latent variables; $SDEP_{EXT}$: standard deviation error of prediction (or root mean squared error of prediction, RMSEP) for the external test set

P	$PC_{FL:SL}$	$SDEP_{EXT}$
A	1:3	0.89

Multi-Probe Guided Region Selection Test Set predictions: SDEP values considering the optimal first level and second level PCs. P:AutoGrid Multi-Probe; $PC_{FL:SL}$: optimal first level and second level PC; $SDEP_{EXT}$: standard deviation error of prediction (or root mean squared error of prediction, RMSEP) for the external test set





HIT-RATES

RANK_BY_NUMBER	HIT_RATES 6 (10%)		HIT_RATES 12 (20%)	
QSAR, and MEAN (A, NA, HD, MP)	3.00	(50.00%)	6.00	(50.00%)
QSAR and MP	3.00	(50.00%)	6.00	(50.00%)

➢ Anti-TB case study

- ✓ Quantitative pharmacophore model
- ✓ Useful conformational results to design new compounds
- ✓ Capability to correctly predict low active compounds
- ✓ Capability to discriminate most and least active compounds (CHEMBL)
- ✓ 120 compounds were identified from the NCI Diversity Set as new potential anti-TB agents

European Journal of Medicinal Chemistry 63 (2013) 765–781

Contents lists available at SciVerse ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Original article

Design, synthesis and biological evaluation of new classes of thieno [3,2-*d*]pyrimidinone and thieno[1,2,3]triazine as inhibitor of vascular endothelial growth factor receptor-2 (VEGFR-2)

Enrico Perspicace [a], Valérie Jouan-Hureaux [d], Rino Ragno [b,*], Flavio Ballante [b], Stefania Sartini [c], Concettina La Motta [c], Federico Da Settimo [c], Binbin Chen [a], Gilbert Kirsch [a], Serge Schneider [e], Béatrice Faivre [d], Stéphanie Hesse [a,**]

[a] *Laboratoire d'Ingénierie Moléculaire et Biochimie Pharmacologique, UMR CNRS 7565 SRSMC, Institut Jean Barriol, FR CNRS 2843, Université de Lorraine, 1 Boulevard Arago, 57070 Metz, France*
[b] *Rome Center for Molecular Design, Dipartimento di Chimica e Tecnologie del Farmaco, Università degli Studi di Roma "La Sapienza", P. le A. Moro 5, 00185 Roma, Italy*
[c] *Dipartimento di Farmacia, Università di Pisa, Via Bonanno 6, 56126 Pisa, Italy*
[d] *EA 4421 SiGReTO, Université de Lorraine, Faculté de pharmacie, 5-7 Rue Albert Lebrun, BP80403, 54001 Nancy cedex, France*
[e] *Laboratoire National de Santé, Division de Toxicologie, Université de Luxembourg, 162a Avenue de la Faïencerie, L-1511 Luxembourg, Luxembourg*

ARTICLE INFO

Article history:
Received 13 December 2012
Received in revised form
7 March 2013
Accepted 10 March 2013
Available online 19 March 2013

Keywords:
Vascular endothelial growth factor receptor-2 (VEGFR-2)
Anti-angiogenic activity
Structure-based drug design (SBDD)
Ligand-based drug design (LBDD)
3-D QSAR
Thieno[3,2-*d*]pyrimidinone
Thieno[1,2,3]triazine
Endothelial cell tube formation

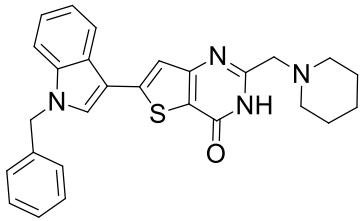
ABSTRACT

Driven by a multidisciplinary approach combination (Structure-Based (SB) Three-Dimensional Quantitative Structure–Activity Relationships (3-D QSAR), molecular modeling, organic chemistry and various biological evaluations) here is reported the disclosure of new thienopyrimidines **1**–**3** as inhibitors of KDR activity and human umbilical vein endothelial cell (HUVEC) proliferation. More specifically, compound **2f** represents a new lead compound that inhibits VEGFR-2 and HUVEC at μM concentration. Moreover by the mean of an endothelial cell tube formation *in vitro* model **2f** tartaric acid salt proved to block angiogenesis of HUVEC at μM level.

- 3D-QSAR models and pIC_{50} predictions of a set of **55** new compounds ($5.82<pIC_{50}<7.82$)
- Experimental activity (%inhibition @ 200µM) evaluated for **51** compounds.(ranging from 13,7% to 99.2%. (4 inactive compounds)

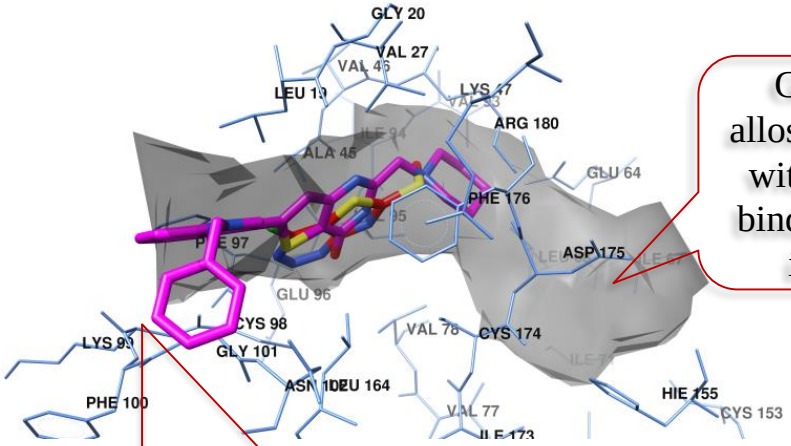
2-indolyl-thieno[3,2-*d*]pyrimidinone

2f



99.2% inhibition at 200µM
IC_{50}=2.25±0.1 µM (pIC_{50}=5,54)
Predicted **pIC_{50}= 8.18**

Docked binding modes in the ATP binding site of VEGFR-2



Gap in the allosteric pocket with potential binding for the inhibitor

How to improve this activity ?

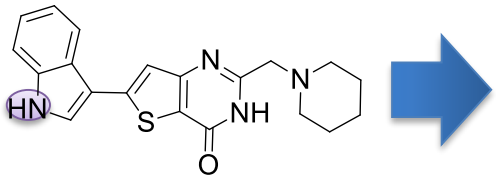
Retention of the compound outside of the ATP binding site

Design of potent VEGFR-2 inhibitors: 3-D QSAR considerations

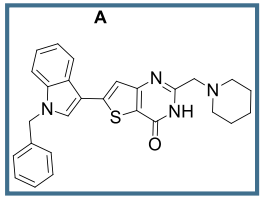
2 modulations of A were studied.
Prediction of pIC_{50}.

1. Protection of the indole moiety

2. Isomery of the thienopyrimidinone

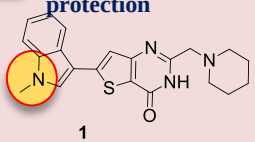


Unprotected indole = **no activity**!



pIC_{50}=8.18

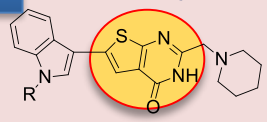
1 **Indole protection**



pIC_{50}=6.81

Decreased inhibitory activity compared to A

2 **Isomery**



pIC_{50}=8.37 for **2**

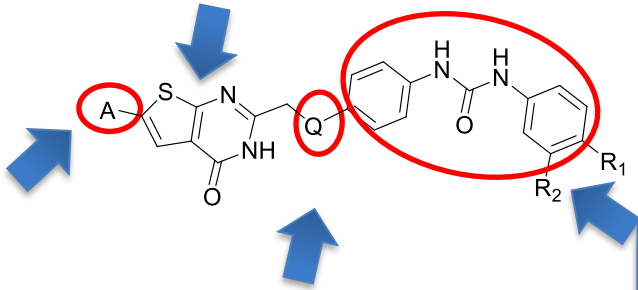
pIC_{50}=8.51 for **3**

Increased inhibitory activity compared to A

Lead Optimization

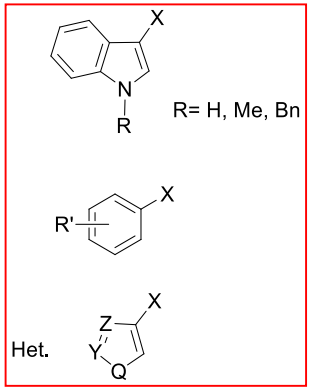
3 potential modulations

Thieno[2,3-*d*]pyrimidinone isomer



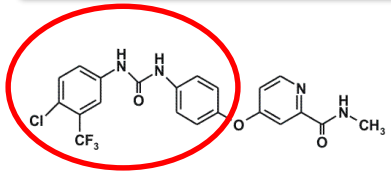
R_1,R_2=H
R_1=F, R_2=H
R_1=Cl, R_2=CF_3

1 **A= I, Br**
Introduction of aromatic moieties



2 **Q= S, O, NH**
Modulation of the heteroatom:
= various possibilities of interaction with the Kinase Domain (KDR)

3 **Phenyl urea:**
extending chain commonly studied in VEGFR-2 inhibitors



Nexavar (Sorafenib) ®

GM. Keating, A. Santoro. Sorafenib: a review of its use in advanced hepatocellular carcinoma. *Drugs* 69 (**2009**) (2): 223–240

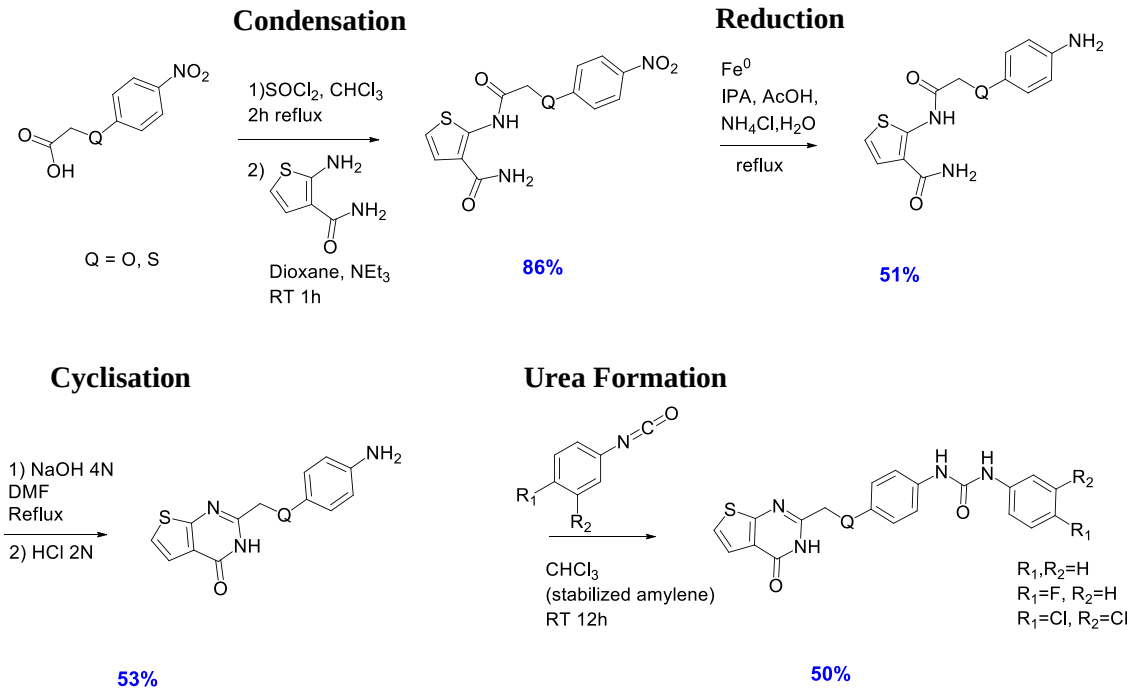
RESEARCH PERIOD ABROAD
(09/2012-03/2013)

Laboratoire d'Ingénierie Moléculaire
LIMBP
et Biochimie Pharmacologique

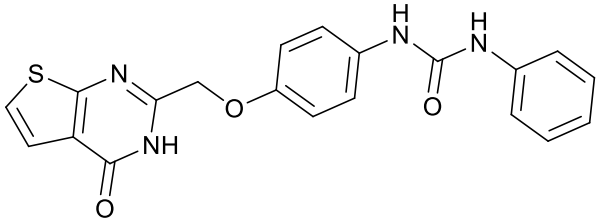
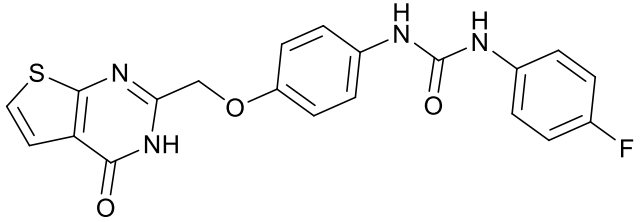
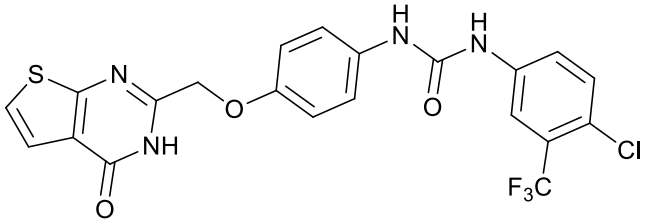
Synthesis of new classes of thienopyrimidinone as Inhibitor of Vascular Endothelial Growth Factor Receptor-2 (VEGFR-2)

Alternative synthetic strategy: access to the unsubstituted thieno[2,3-*d*]pyrimidinones

Linear open pathway: 4 steps synthesis

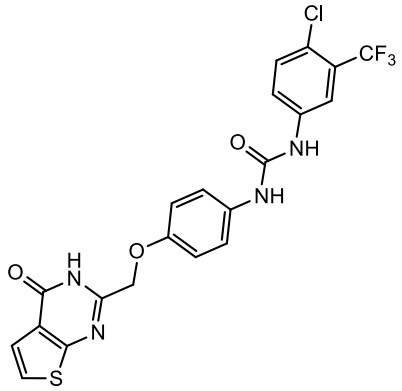


Results of the 3D-QSAR modeling: the 10 best predictions (over 49)

ID	Structure	Predictions	
		pIC_{50}	IC_{50} nM
7a		7.73	18.72
7b		7.63	23.47
7c		7.57	26.74

1-[4-chloro-3-(trifluoromethyl)phenyl]-3-[4-({4-oxo-3H-thieno[2,3-d]pyrimidin-2-yl}methoxy)phenyl]urea

Molecular Weight	494.874
Aspect	Light grey solid



^{1}H NMR (400MHz, DMSO): δ_H 12.57(bs, 1H), 9.07(s, 1H), 8.65(s, 1H), 8.09(s, 1H), 7.64-7.58(m, 3H), 7.41-7.38(m, 3H), 7.01-6.99(m, 2H), 5.00(s, 2H)

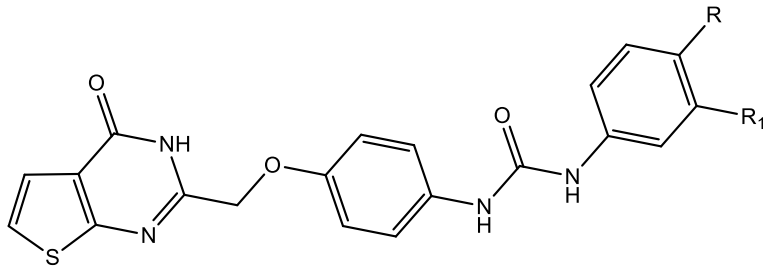
^{19}F NMR (376 MHz, DMSO): δ_F -61.44(s, 3F)

MS (ESI, $[M+Na]^+$) *m/z* for C21H14ClF3N4NaO3S	theoretic: 517.0325 obtained: 517.0323

		C	H	N	S
Microanalysis:	theoretic:	50.97	2.85	11.32	6.48
	obtained:	51.08	2.77	11.56	6.41

Biological evaluation* for compounds 7a, 7b, 7c

Structures and either EGFR, VEGFR-2 inhibitory activity of thieno[2,3-d]pyrimidinones **7a-7c**



7

#	R	R_1	% Inhibition @ 10 µM EGFR	% Inhibition @ 10 µM VEGFR-2
7a	H	H	n.a.[a]	9%
7b	F	H	n.a.[a]	5%
7c	Cl	CF_3	n.a.[a]	13%

[a] Not active. No inhibition was observed up to 10 µM of the tested compound.

*Depreux Lab.

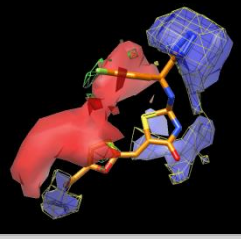
➢ VEGFR-2 case study

✓ 3-D QSARs correctly predicted **2f** as new potent VEGFR-2 inhibitor

✓ Binding mode analysis led to design new compounds

✓ I practiced on organic chemistry

✓ Finally 3 new compounds were synthetised and biologically evaluated



Hsp90 Inhibitors (I). Definition of 3-D QSAutogrid/R Models as a Tool for Virtual Screening

Flavio Ballante, Antonia Caroli, Richard B. Wickersham III And Rino Ragno

Journal of Chemical Information and Modeling,

(accepted with minor revisions, under correction)

Hsp90 Inhibitors (II). Combining ligand-based and structure-based approaches for Virtual Screening application

Antonia Caroli, Flavio Ballante, Richard B. Wickersham III, Federico Corelli And Rino Ragno

Journal of Chemical Information and Modeling,

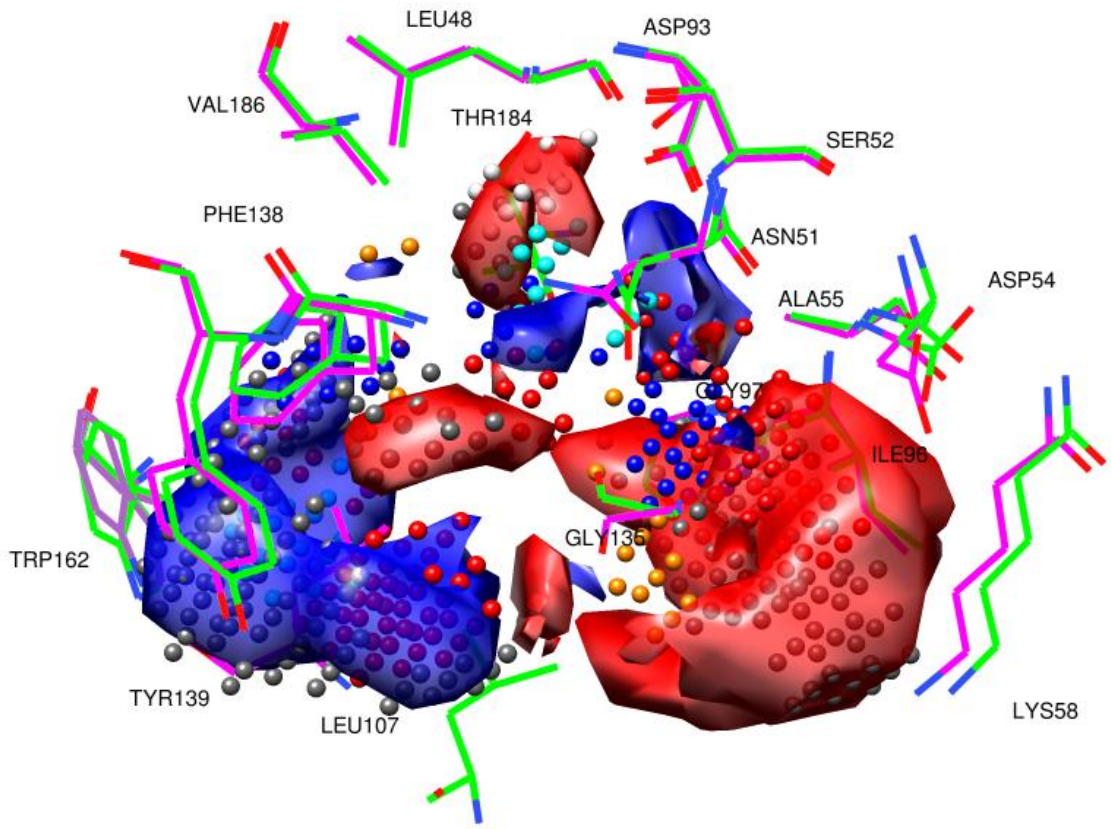
(accepted with minor revisions, under correction)

Autogrid/R PLS models statistical results (CAPP process was applied).

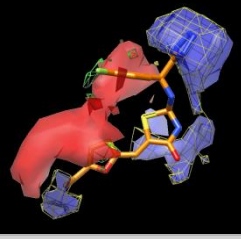
model	P	PC	r^2	q^2_{LOO}	q^2_{K5FCV}	r^2_{YS}	q^2_{YS}
1	A	2	0.93	0.61	0.59	0.69	-0.46
2	C	2	0.93	0.61	0.58	0.69	-0.50
3	HD	2	0.85	0.55	0.54	0.44	-0.52
4	NA	2	0.93	0.61	0.59	0.71	-0.42
5	N	2	0.93	0.62	0.58	0.68	-0.52
6	OA	2	0.93	0.61	0.59	0.69	-0.47
7	e	2	0.93	0.63	0.60	0.73	-0.50
8	d	1	0.72	0.61	0.60	0.12	-0.22

Statistical Results Obtained from MPGRS Analysis

MPGRS 3-D QSAR							
$PC_{FL:SL}$	r^2	q^2_{LOO}	q^2_{K5FCV}	$SDEP_{LOO}$	$SDEP_{K5FCV}$	r^2_{YS}	q^2_{YS}
1:2	0.94	0.69	0.68	0.43	0.44	0.67	-0.50



MPGRS. : key points with PLS-coefficients contour maps (contour levels 75%: positive: red; negative: blue). 3OWD in magenta and 1UYD in green. The points are color coded according to that reported in the methodology reference.



Test Set predictions

P	PC	$SDEP_{EXT}$
A	2	0.79
C	2	0.79
HD	2	0.78
NA	2	0.80
N	2	0.79
OA	2	0.81
e	2	0.79
d	1	0.86

MPGRS. Multi Probe model Test Set predictions.

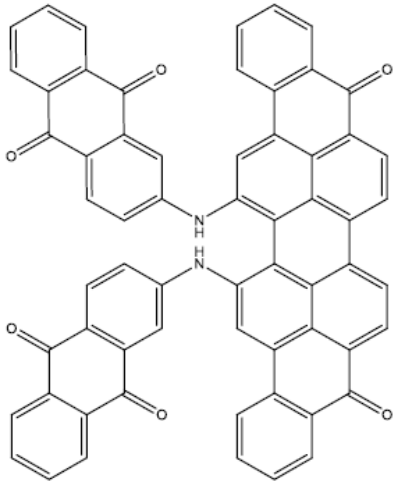
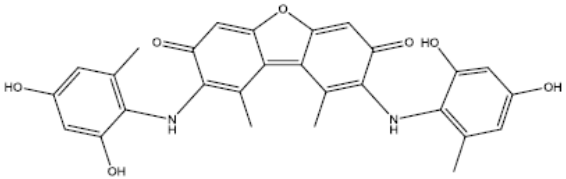
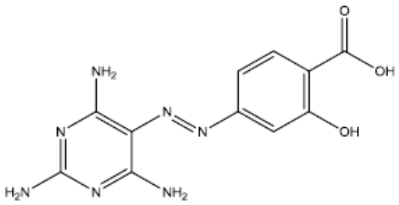
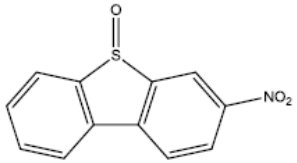
P	$PC_{EL:SL}$	$SDEP_{EXT}$
AutoGrid MP	1:2	0.81

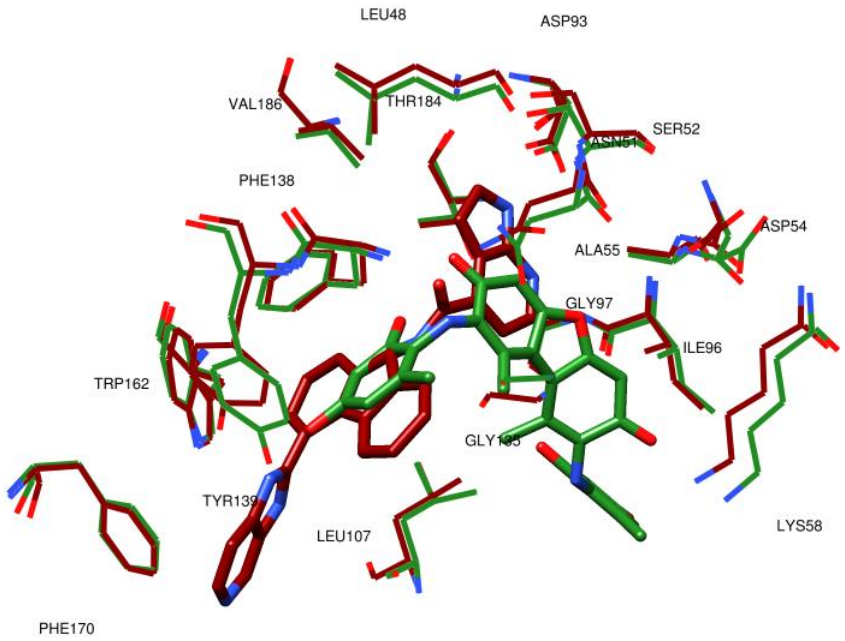
VS

1785 compounds (NCI Diversity Set)

80 molecules selected for biological assays

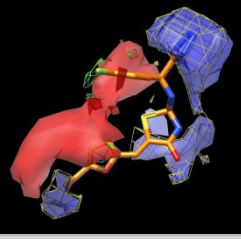
4 derivatives with IC_{50} values ranging between 18-63 μM

	ID	IC_{50} [μM]
	**NCI23128**	18
	**NCI610930**	32
	**NCI117285**	49
	**NCI170578**	63



The most active screened compound **NCI610930 (**green coloured) in its BC system (protein and pose) overlapped with 2YKI[1].

[1]J Med Chem 2011, 54, 7206-19.



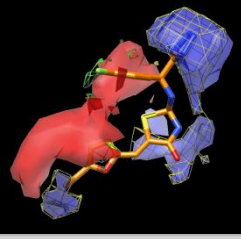
➢ HSP90 case study

✓ Max usage of available HSP90 crystal structures

✓ 3-D QSARs allowed to derive several LB-SB convergence points

✓ An optimized VS procedure was applied on the NCI Diversity Set

✓ 4 compounds were identified to be active in a low micromolar range



COMBINE:
Comparative Binding Energy Analysis

combination of data

(structures of the receptor-ligand
the measured biological activity)

(combination of molecular mechanics
and chemometrics for the analysis)

While the performed chemometric analysis is similar to CoMFA, the analyzed data, in COMBINE, include the interaction energies between the receptor and the ligand rather than the only interaction ligands properties.

Rebecca C. Wade, Angel R. Qrtiz, e Federico Gago, «**Comparative Binding Energy Analysis**», in *3D QSAR in Drug Design*, cur da. Hugo Kubinyi, Gerd Folkers, e Yvonne C. Martin, vol. 2 (Dordrecht: Kluwer Academic Publishers, 2002), 19-34, http://www.springerlink.com/content/r1g65852226264h3/.

The COMBINE models allow to highlight the fundamental interactions that drive for the different binding mode characterizing the training set.

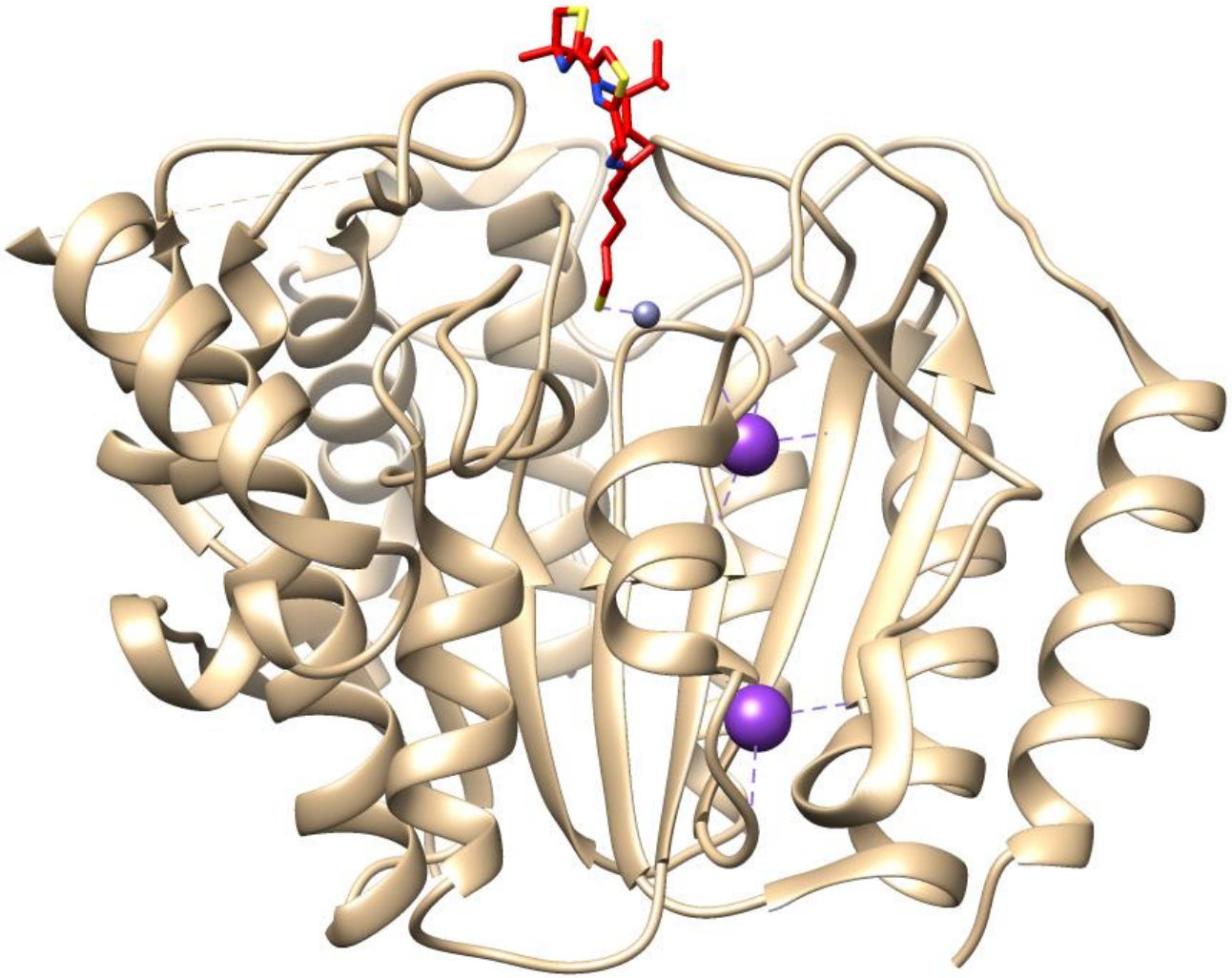
The used training set can be characterized by:

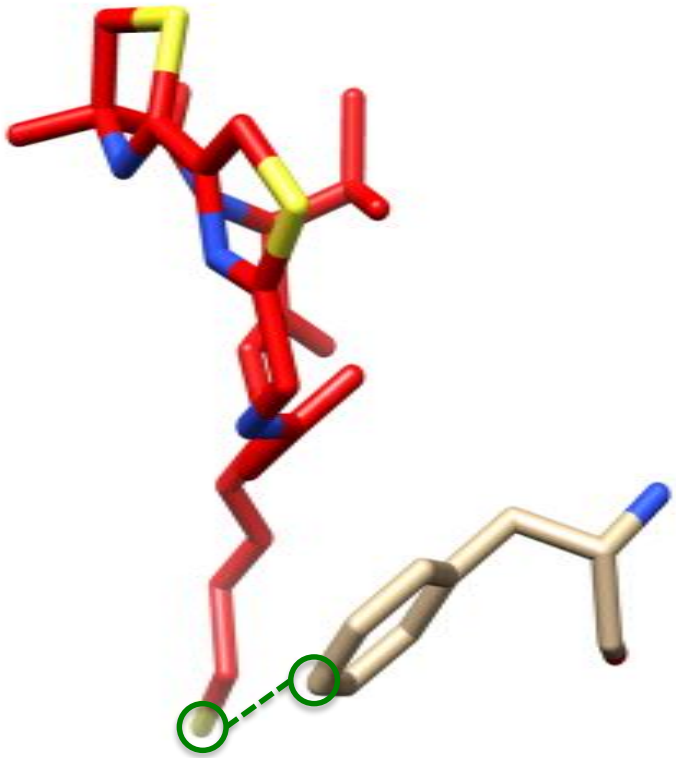
✓ complexes consisting of the same receptor complexed with different ligands

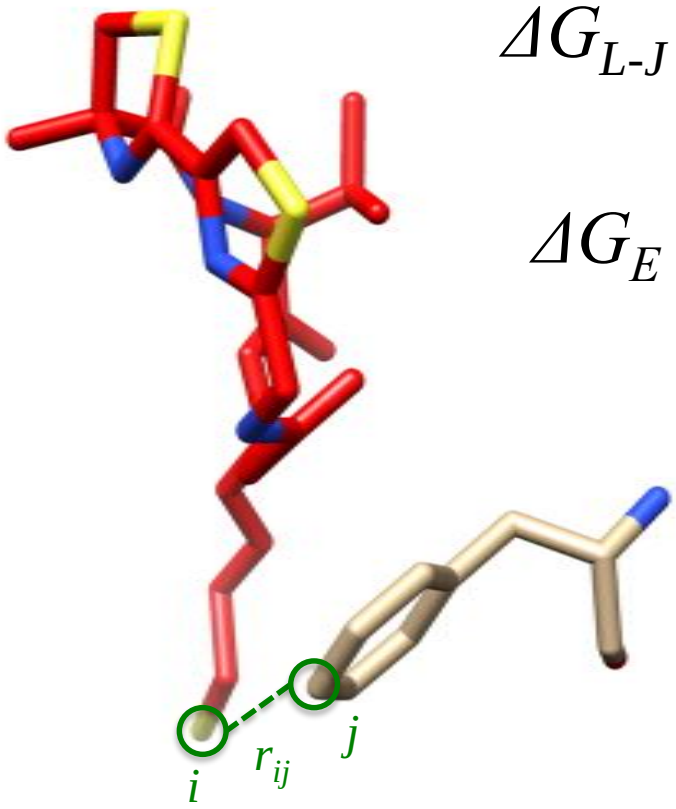
✓ different receptors complexed with the same ligand

✓ different ligands complexed with different receptors

➢ **A simplified and faster version of comparative binding energy (COMBINE) analysis, named COMBINEr was developed…**

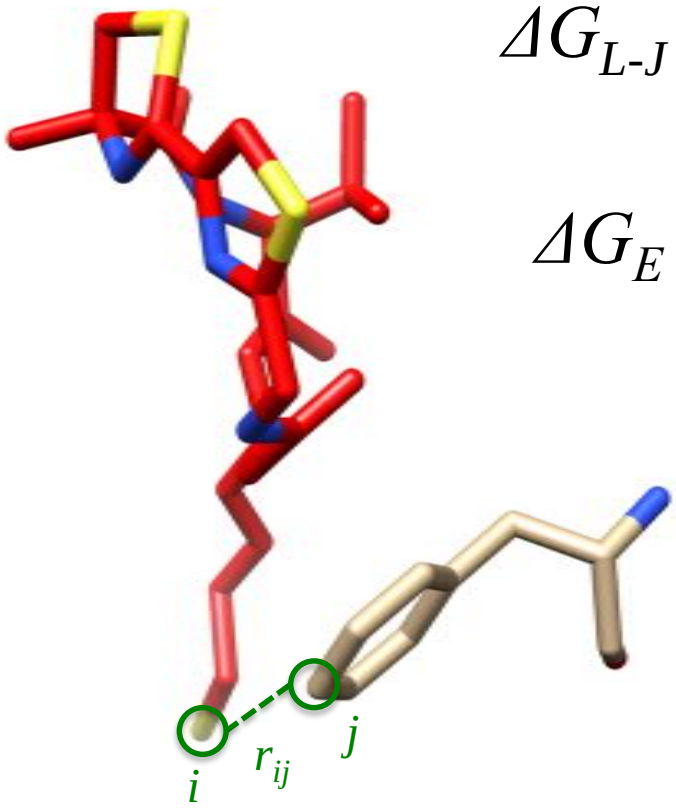






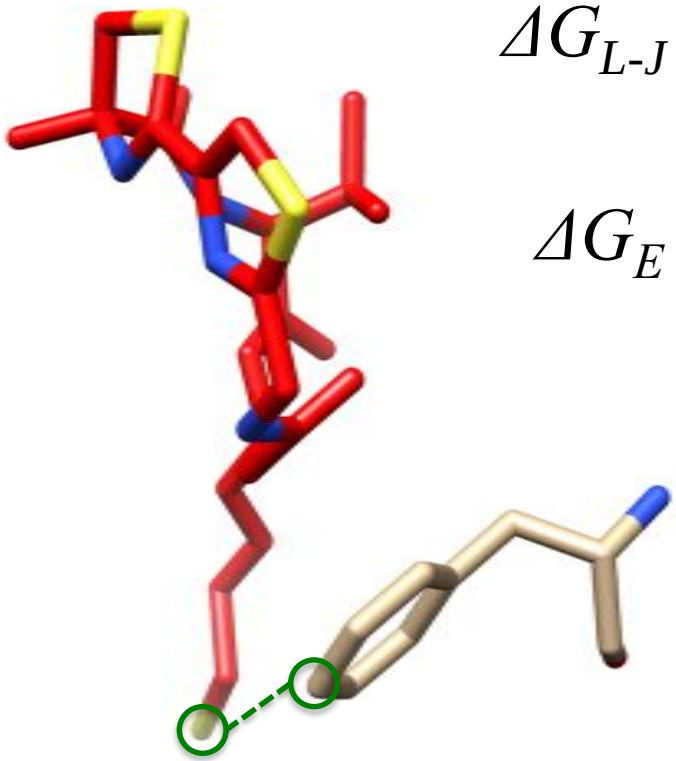
$$\Delta G_{L\text{-}J} = \varepsilon \left[\left(\frac{r_m}{r_{ij}} \right)^{12} - 2 \left(\frac{r_m}{r_{ij}} \right)^{6} \right]$$

$$\Delta G_E = \frac{1}{4\pi\varepsilon_0} \frac{q_1 q_2}{r_{ij}}$$



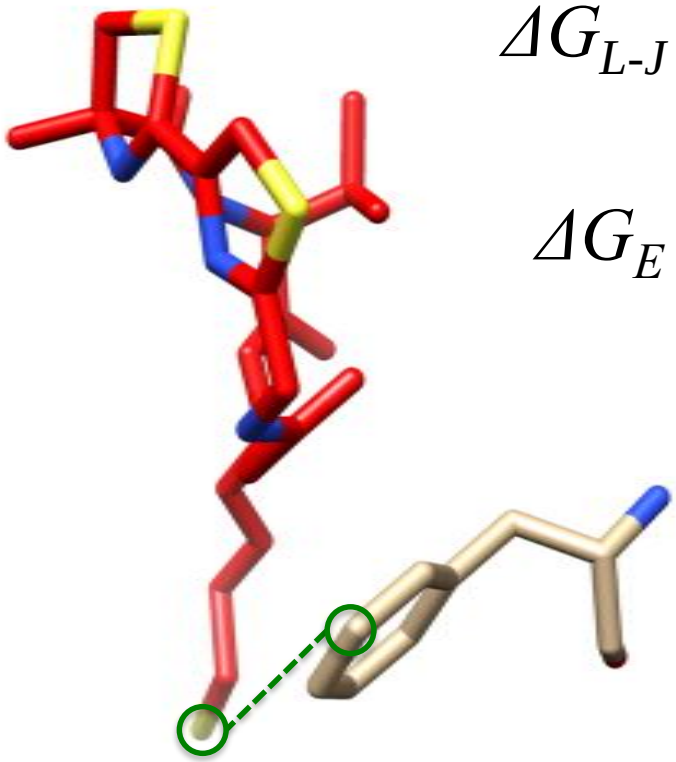
$$\Delta G_{L\text{-}J} = \sum_{ij} \varepsilon \left[\left(\frac{r_m}{r_{ij}} \right)^{12} - 2 \left(\frac{r_m}{r_{ij}} \right)^{6} \right]$$

$$\Delta G_E = \sum_{ij} \frac{1}{4\pi\varepsilon_0} \frac{q_1 q_2}{r_{ij}}$$



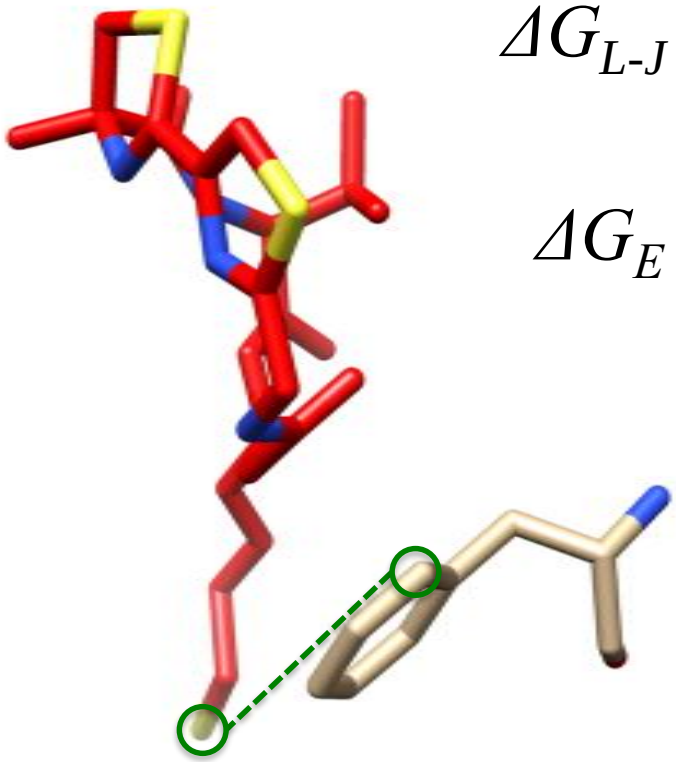
$$\Delta G_{L\text{-}J} = -0.356$$

$$\Delta G_{E} = -0.849$$



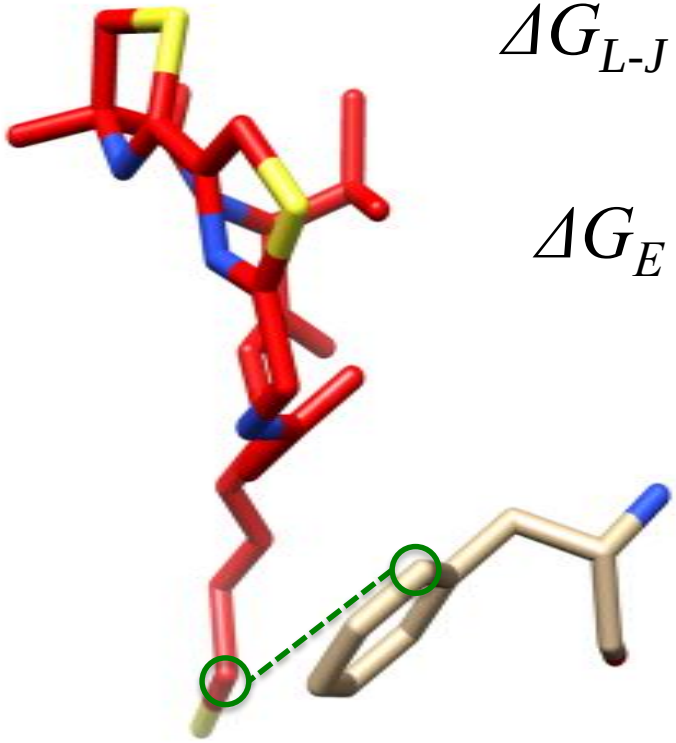
$$\Delta G_{L\text{-}J} = -0.356+(-1.287)$$

$$\Delta G_{E} = -0.849+(-0.570)$$



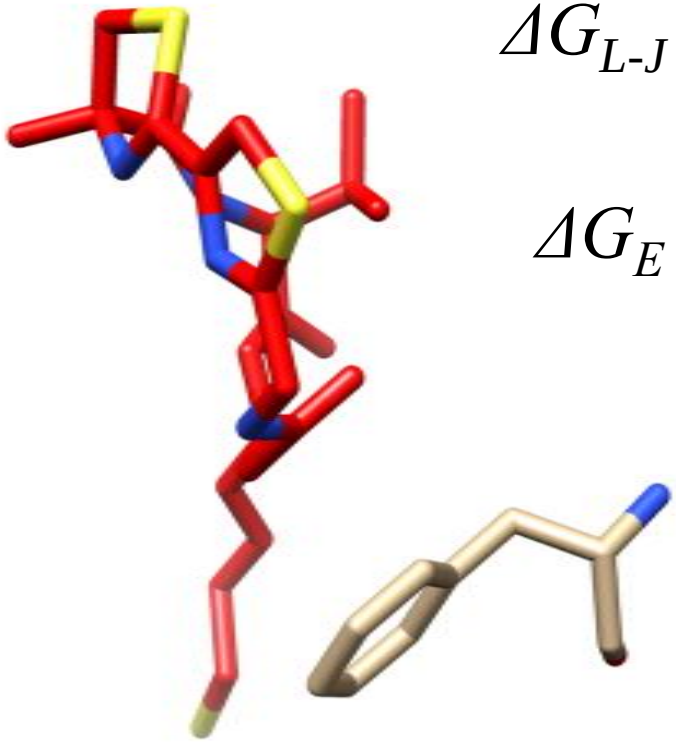
$$\Delta G_{L\text{-}J} = -0.356+(-1.287)+(-0.749)+\ldots$$

$$\Delta G_{E} = -0.849+(-0.570)+(-1.136)+\ldots$$



$$\Delta G_{L\text{-}J} = -0.356+(-1.287)+(-0.749)+\ldots$$

$$\Delta G_{E} = -0.849+(-0.570)+(-1.136)+\ldots$$

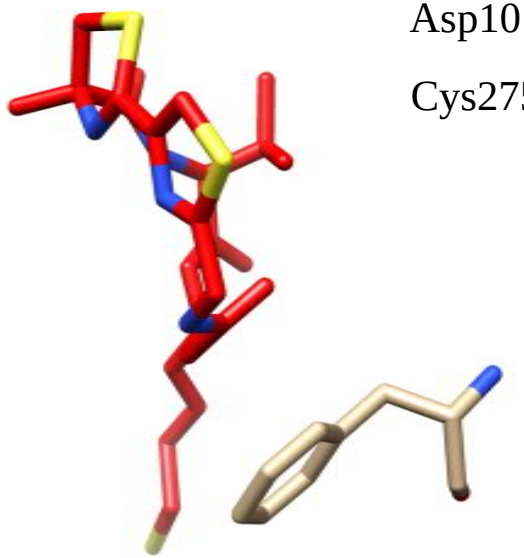


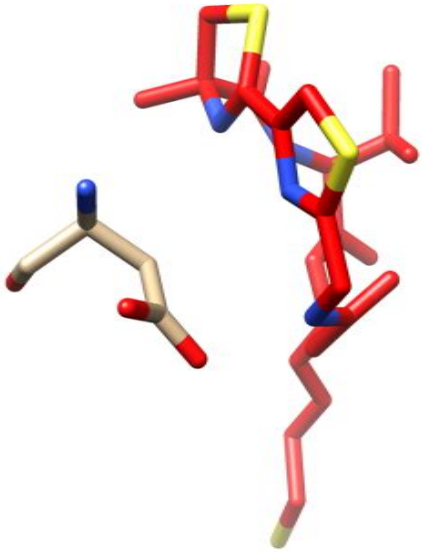
$$\Delta G_{L\text{-}J} = -6.725$$

$$\Delta G_{E} = -8.430$$

$\Delta G_{L\text{-}J}$	compd1	compd2	compd3	compd4
Phe208	1.236			
Asp101				
Cys275				

ΔG_{E}	compd1	compd2	compd3	compd4
Phe208	2.851			
Asp101				
Cys275				



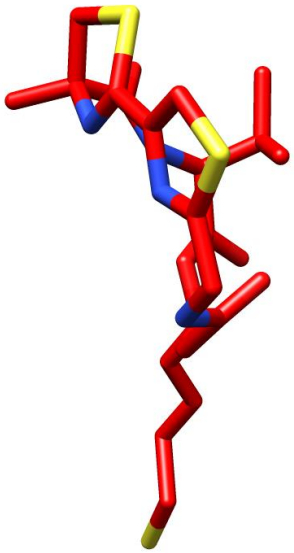


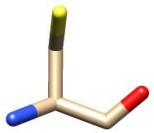
$\Delta G_{L\text{-}J}$	compd1	compd2	compd3	compd4
Phe208	1.236			
Asp101	-6.725			
Cys275				

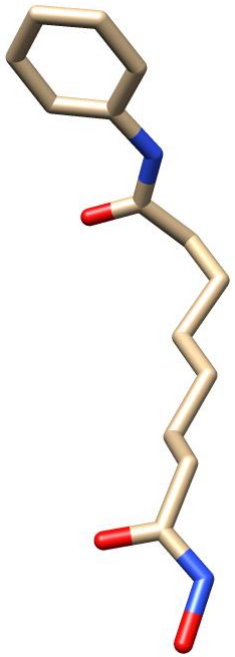
ΔG_{E}	compd1	compd2	compd3	compd4
Phe208	2.851			
Asp101	-8.430			
Cys275				

$\Delta G_{L\text{-}J}$	compd1	compd2	compd3	compd4
Phe208	1.236			
Asp101	-6.725			
Cys275	-6.246			

ΔG_{E}	compd1	compd2	compd3	compd4
Phe208	2.851			
Asp101	-8.430			
Cys275	-7.879			

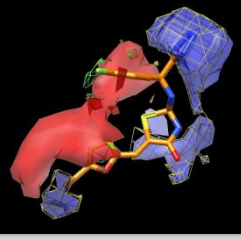






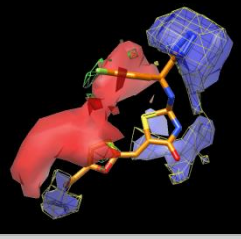
$\Delta G_{L\text{-}J}$	compd1	compd2	compd3	compd4
Phe208	1.236	-0.489	…	…
Asp101	-6.725	-5.213	…	…
Cys275	-6.246	-3.551	…	…

ΔG_{E}	compd1	compd2	compd3	compd4
Phe208	2.851	-6.187	…	…
Asp101	-8.430	-7.488	…	…
Cys275	-7.879	-5.375	…	…



V_{L-J}	compd1	compd2	compd3	compd4	
Phe208	$1.236\beta_1$	$-0.489\beta_1$	$\ldots\beta_1$	$\ldots\beta_1$	$\beta_1 X_1$
Asp101	$-6.725\beta_2$	$-5.213\beta_2$	$\ldots\beta_2$	$\ldots\beta_2$	$\beta_2 X_2$
Cys275	$-6.246\beta_3$	$-3.551\beta_3$	$\ldots\beta_3$	$\ldots\beta_3$	$\beta_3 X_3$
.					
.					
.					
IC_{50}	6.3×10^{-7}	3×10^{-9}	4.3×10^{-6}	5.6×10^{-5}	Y

$$\beta_1 X_1 + \beta_1 X_1 + \ldots + \beta_n X_n = Y$$



➢ COMBINEr

✓ Identify the most influent ligand protein interactions

✓ WO patent 61/842,191

J Comput Aided Mol Des (2012) 26:907–919
DOI 10.1007/s10822-012-9586-6

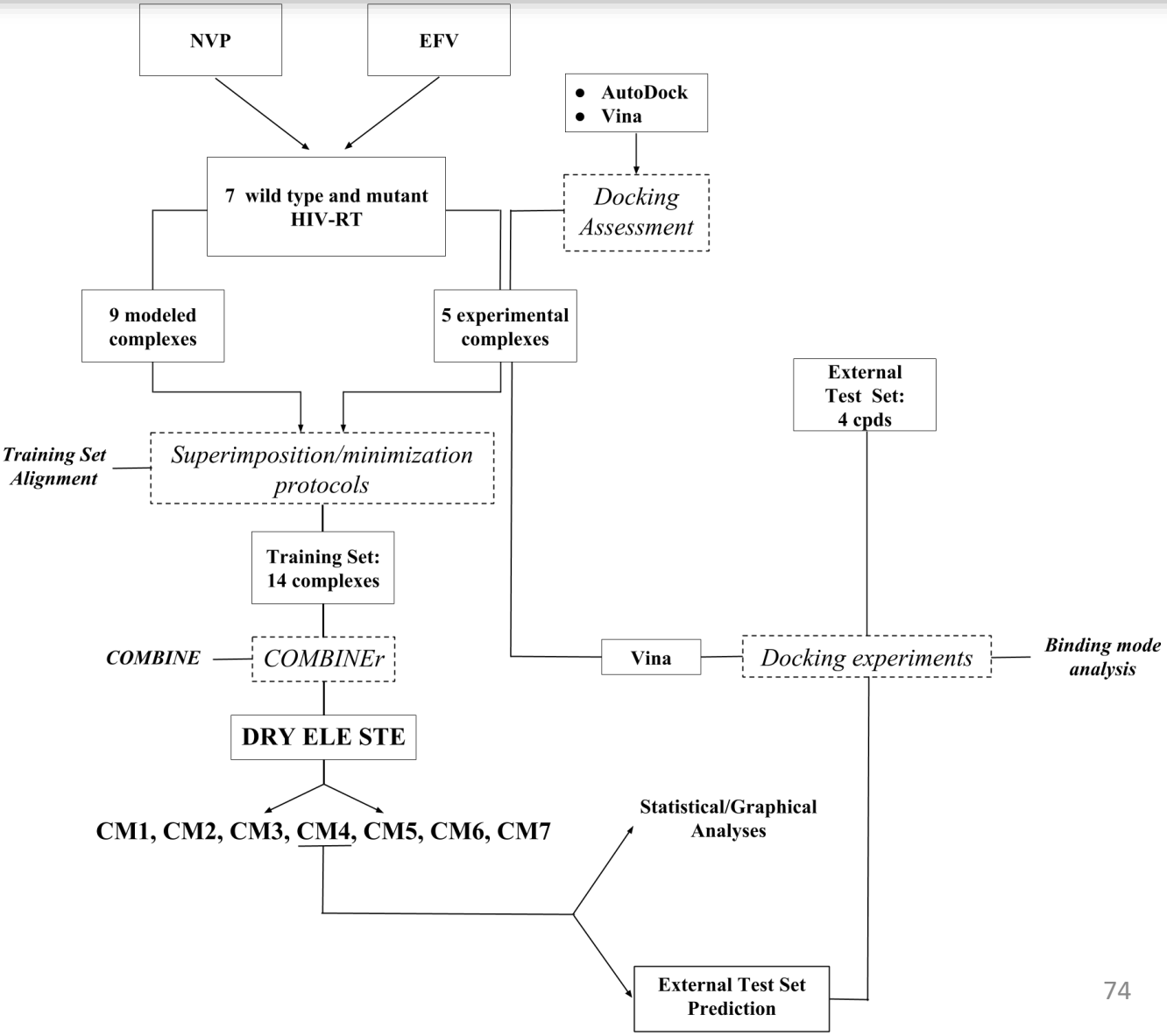
Comprehensive model of wild-type and mutant HIV-1 reverse transciptases

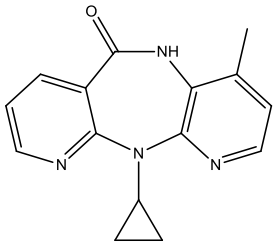
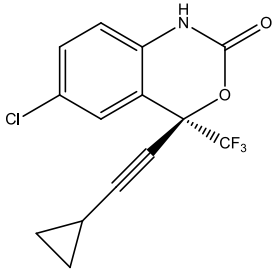
Flavio Ballante · Ira Musmuca · Garland R. Marshall · Rino Ragno

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*[,†]

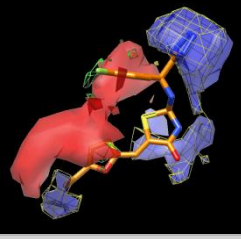


Nevirapine (NVP)	Efavirenz (EFV)
	

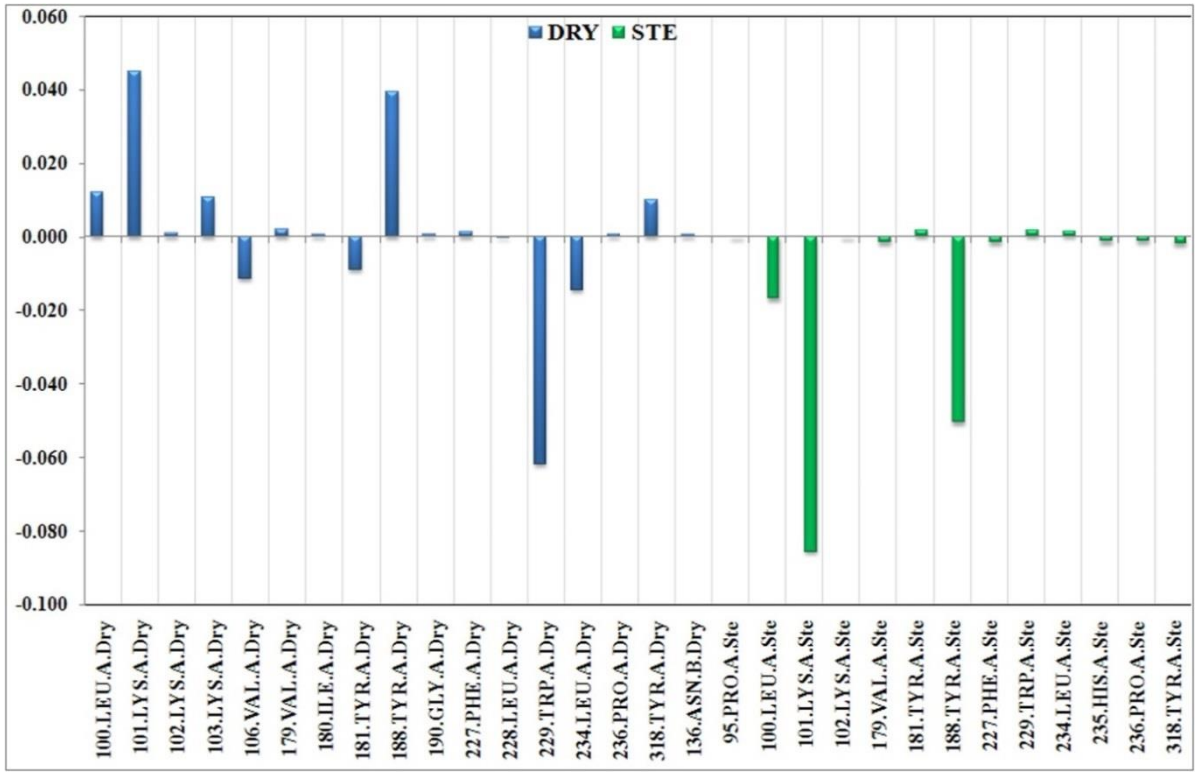
Nevirapine and Efavirenz.

Anti-RT activities (μM) of NVP and EFV used to build the COMBINEr model.

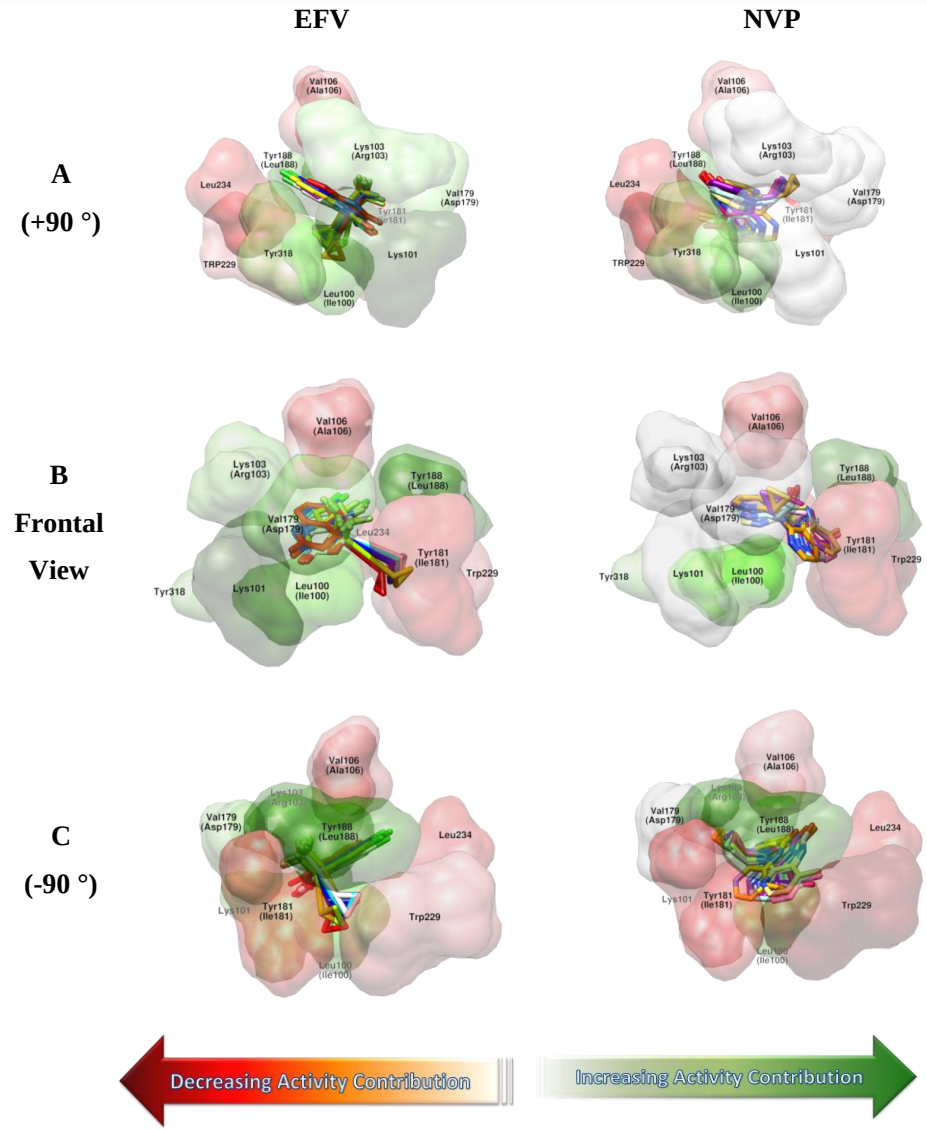
RT	NVP	EFV
WT	0.4	0.03
L100I	9.0	0.12
K103N	7.0	0.16
V106A	10.0	0.04
V179D	2.0	0.10
Y181I	36.0	0.15
Y188L	18.0	0.38

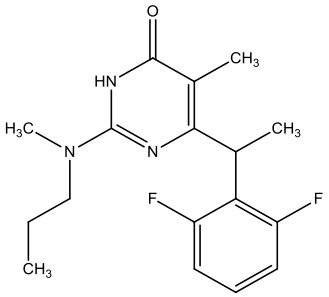
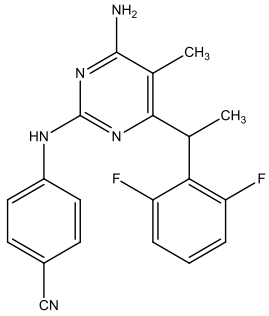


CM	Model	r^2	SDEC	q^2_{LOO}	SDEP $_{LOO}$	q^2_{LSO5}	SDEP $_{LSO5}$	q^2_{LSO2}	SDEP $_{LSO2}$
1	DRY	0.91	0.31	0.82	0.43	0.79	0.46	0.63	0.58
2	ELE	0.80	0.45	0.51	0.71	0.49	0.72	0.37	0.79
3	STE	0.81	0.44	0.69	0.57	0.65	0.60	0.52	0.68
4	DRY_STE	0.88	0.35	0.78	0.48	0.75	0.50	0.61	0.61
5	ELE_STE	0.82	0.43	0.58	0.66	0.53	0.69	0.44	0.75
6	DRY_ELE	0.89	0.34	0.66	0.59	0.63	0.62	0.48	0.70
7	DRY_ELE_STE	0.86	0.38	0.66	0.59	0.62	0.62	0.50	0.70



PLS coefficients x SD obtained from the CM4 model. Only bars with values higher than 0.001 and lower than -0.001 are shown.

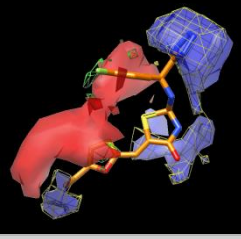


(*R/S*) **MC1501** (*N*,*N*-DABO)	(*R/S*) **MC2082** (DAPY-DABO Hybrid)
	

Racemic DABO derivatives used as external test set.

Experimental and COMBINE model CM4 predicted activities of MC compounds.

	MC1501				**MC2082**			
	R		***S***		***R***		***S***	
	Exp	Pred	Exp	Pred	Exp	Pred	Exp	Pred
WT	8.70	7.46	6.93	7.20	6.81	7.21	4.52	5.77
V106A	8.52	9.19	6.45	5.78	9.52	9.43	6.62	7.51
K103N	7.02	7.17	6.01	7.52	8.52	9.11	7.19	7.52
L100I	7.02	6.69	4.40	7.11	8.10	7.49	6.74	6.03
Y188L	6.71	7.51	4.40	5.11	8.10	7.09	4.40	5.95
Y181I	6.35	6.05	4.40	6.12	6.12	6.25	6.29	5.48



➢ HIV case study

✓ First COMBINEr application

✓ Identify the most influent ligand protein interactions

✓ Predict DABOs derivatives with the correct eudismic ratio

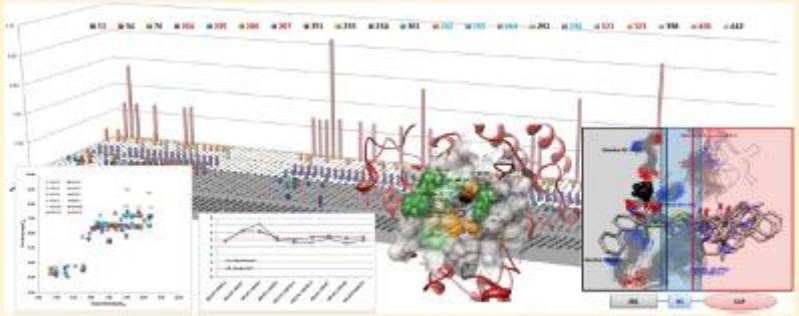
JOURNAL OF CHEMICAL INFORMATION AND MODELING
Article
pubs.acs.org/jcim

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 di Chimica e Tecnologie del Farmaco, Facoltà di Farmacia e Medicina, [‡]Istituto Pasteur—Fondazione Cenci Bolognetti Dipartimento di Chimica e Tecnologie del Farmaco, Facoltà di Farmacia e Medicina, Sapienza Università di Roma, P.le A. Moro 5, 00185 Rome, Italy

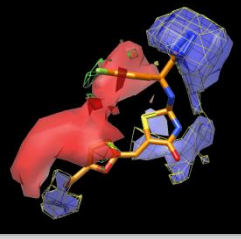
Supporting Information



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.

- Several 3-D QSAR models for the eleven human zinc-based histone deacetylases (HDACs) were derived by a comparative binding energy (COMBINEr) analysis on a series of inhibitors for which biological activities against the 11 human zinc-based HDACs isoforms were available.

- Both ligand-based and structure-based alignments has been used



TRAINING SET:

- 39 complexes derived with crystallized structures

+

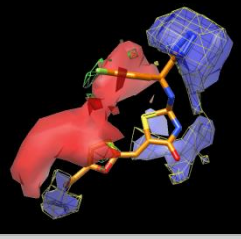
- 55 complexes derived with homology models.

15 INHs

LLX
VALP
NABUT
OXAM
APHA8
SAHA
SBHA
TSA
SCRIP
MS-275
MS-344
NHB
HA3
TFMK

TEST SETS:

- Three different test sets were used for external validation
  - MTS, CTS, LTS

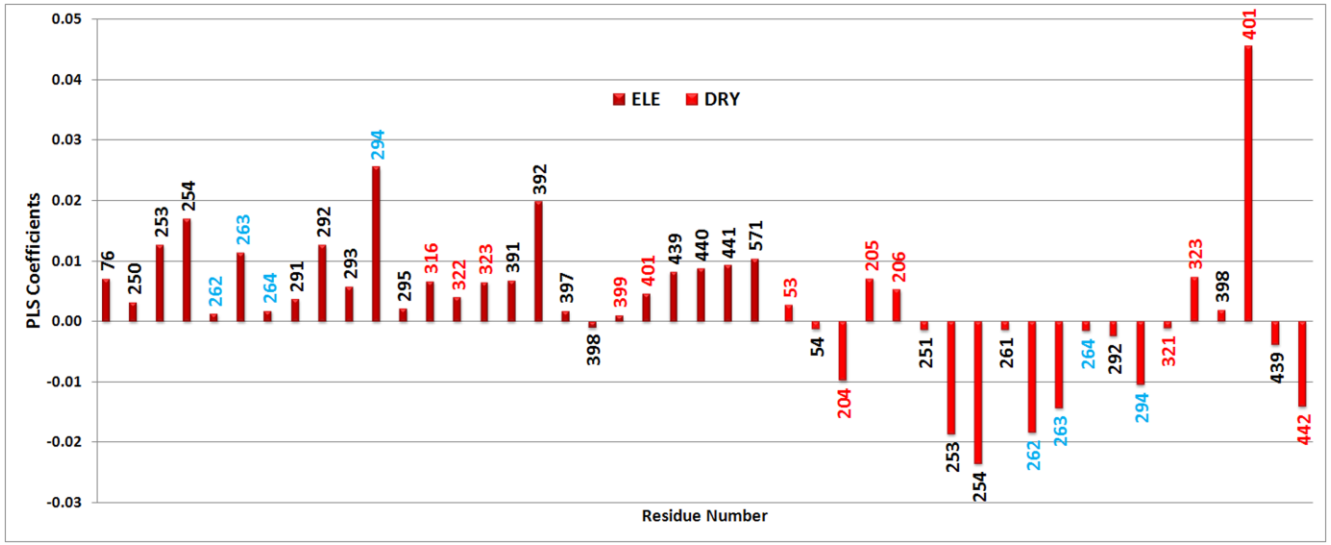


Statistical results of the COMBINE models.

#	Field	PC	r^2	SDEC	q^2_{K5fold}	$SDEP_{K5\ fold}$	q^2_{LOO}	$SDEP_{L\ OO}$	q^2 scrambled % positive values	q^2 scrambled 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

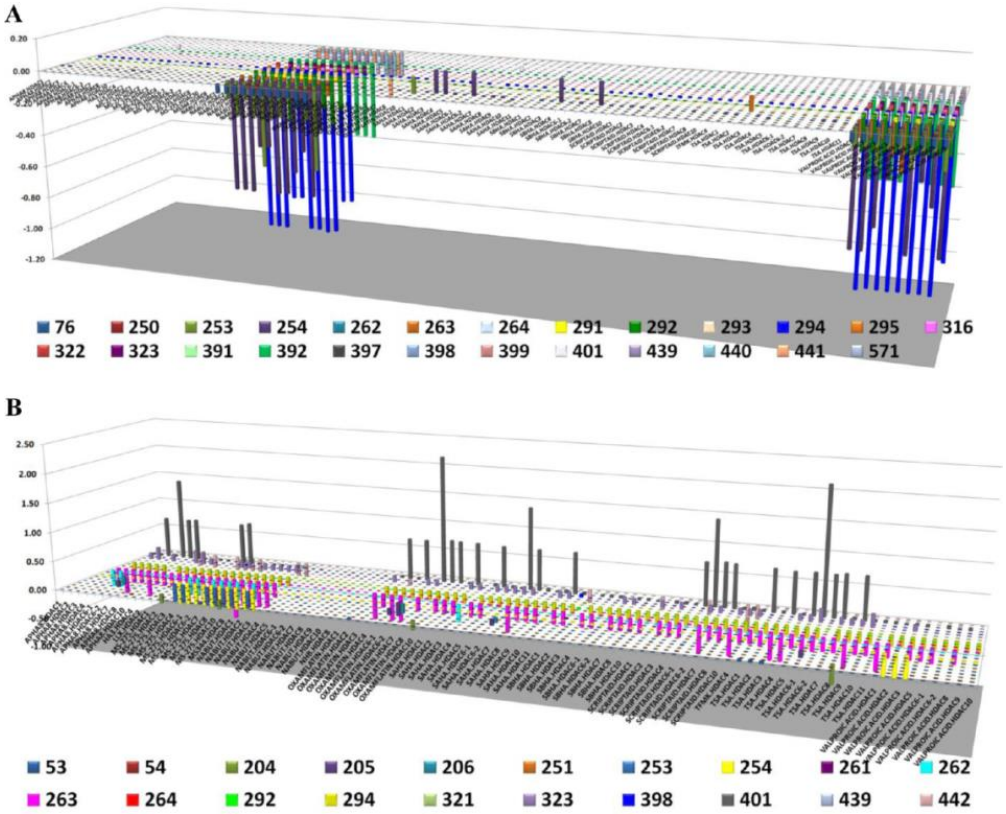
Interaction energies were calculated on a per-residue basis considering electrostatic (ELE), steric (STE) and desolvation (DRY), while statistical PLS-based models were build and validated using in-house R scripts…

…this allowed identification of those residues responsible for both inhibitory activity and selectivity and quantification of their relative importance

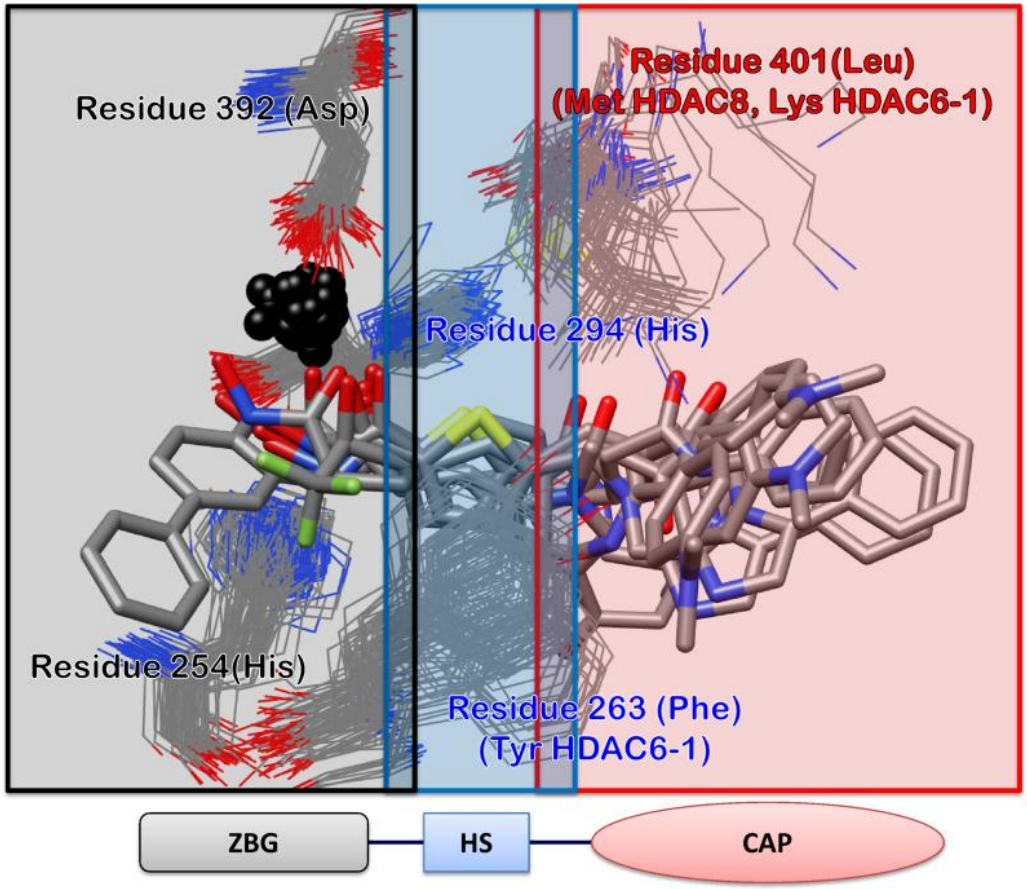


PLS Coeff plot for the ELE+DRY COMBINEr model

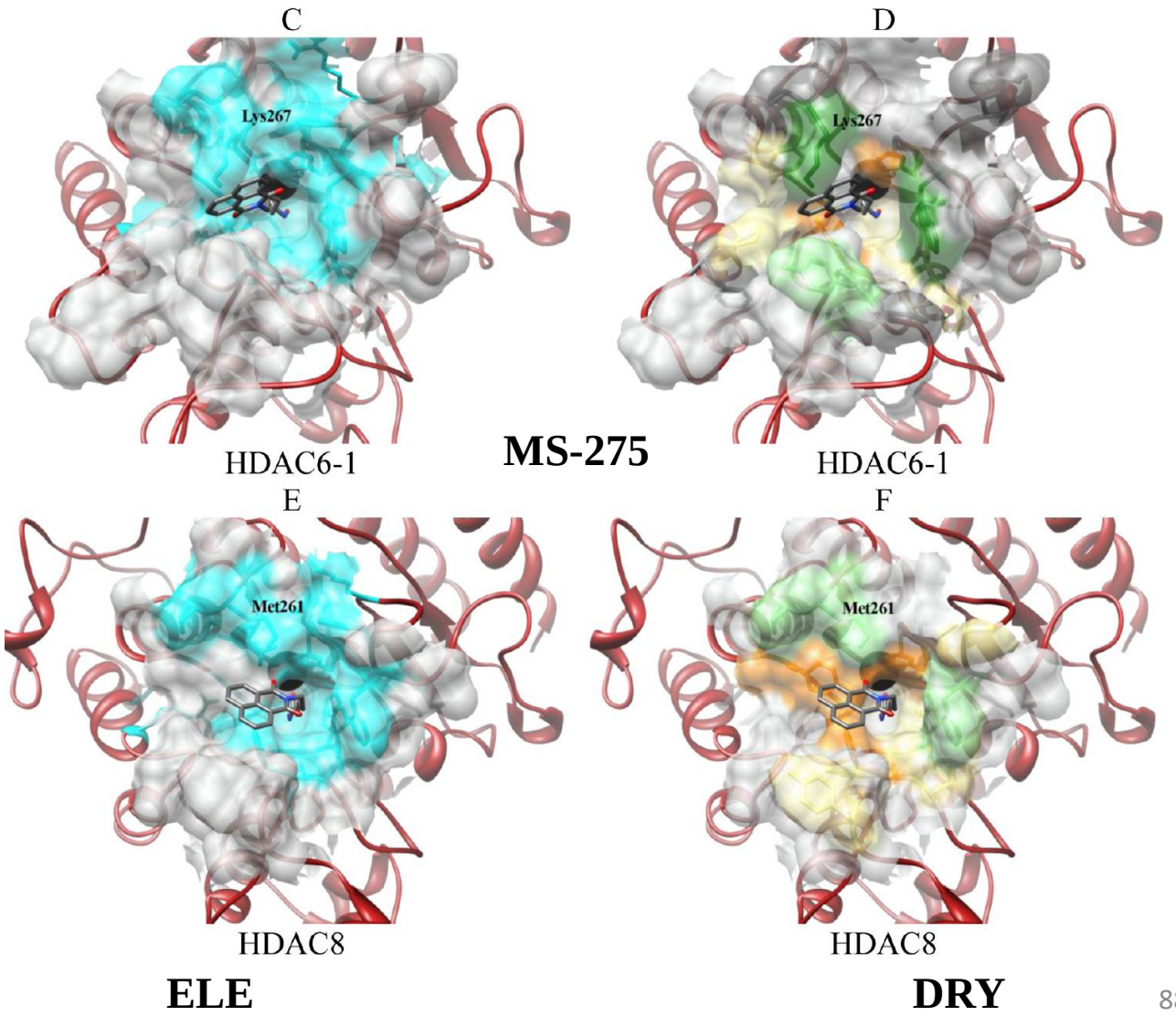
…this allowed identification of those residues responsible for both inhibitory activity and selectivity and quantification of their relative importance

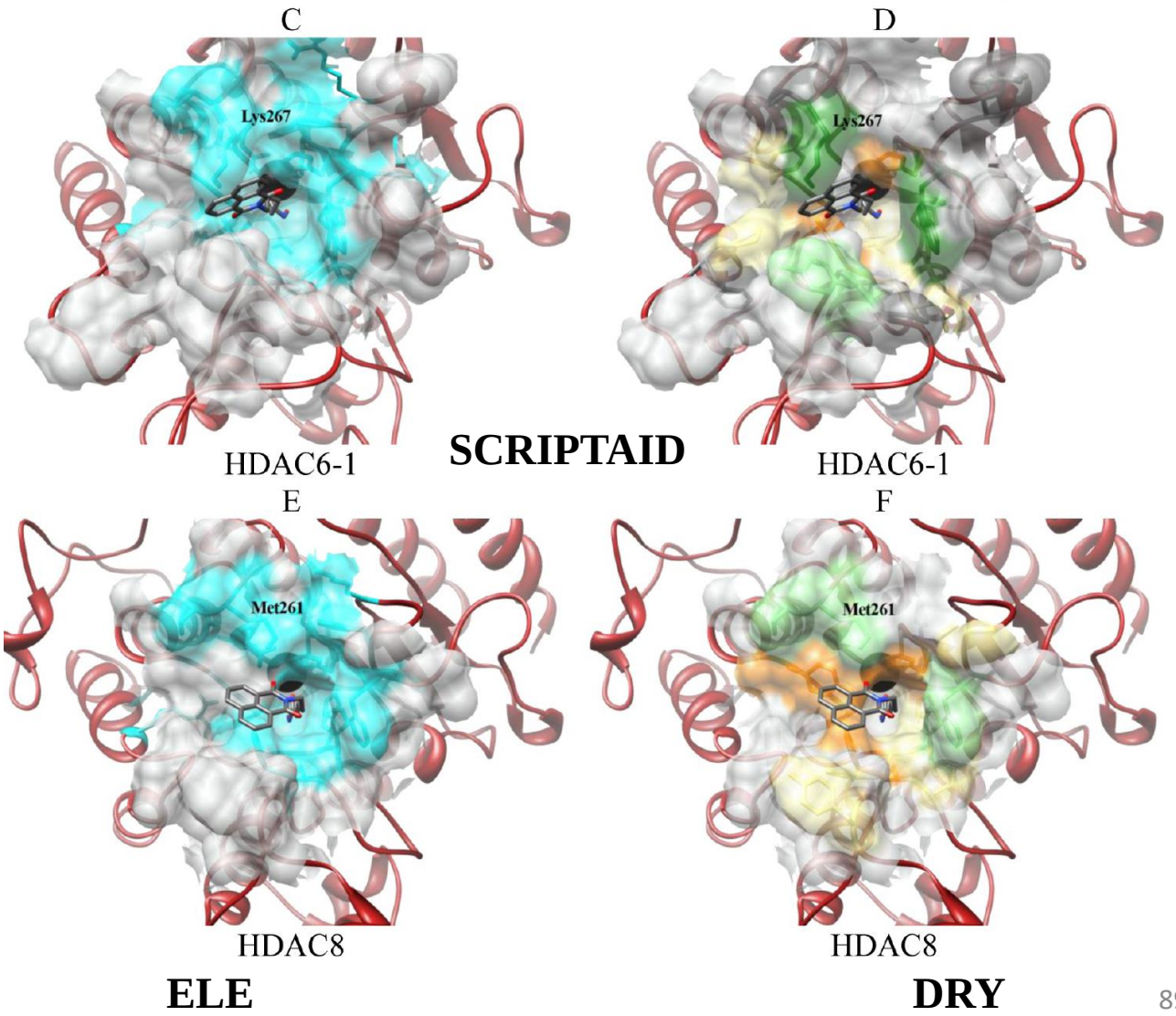


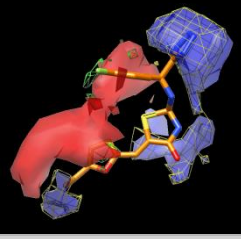
Per-residue activity-contribution plots for the ELE (A) and DRY (B) fields.



The four most important residues (MIRs) from the COMBINE model analysis. The labels and regions are color coded: in red the residues in the HDAC's rim region, in blue those forming the central tube channel and in black those in the proximity of the catalytic Zn ion. ZBG: Zn-binding group, HS: hydrophobic spacer, CAP: hydrophobic capping group.







➢ HDAC case study

✓ The derived COMBINEr model shows good statistical coefficients

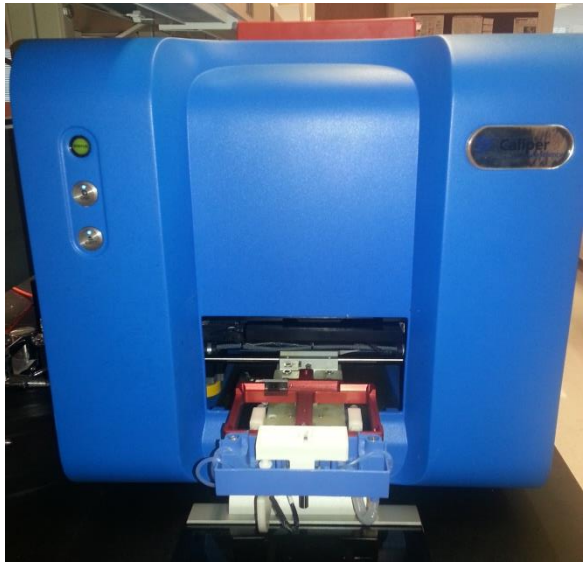
✓ Was predictive for the compounds in the test sets, and robust to cross-validation while omitting multiple data

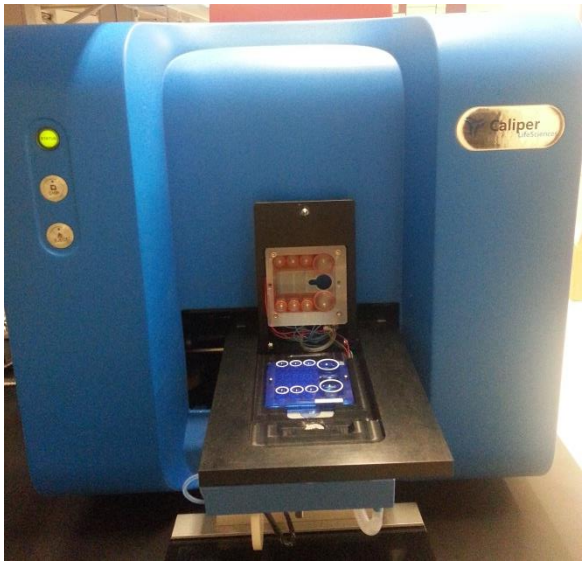
✓ The model was able to rationalize the different activity profiles of the HDAC inhibitors studied

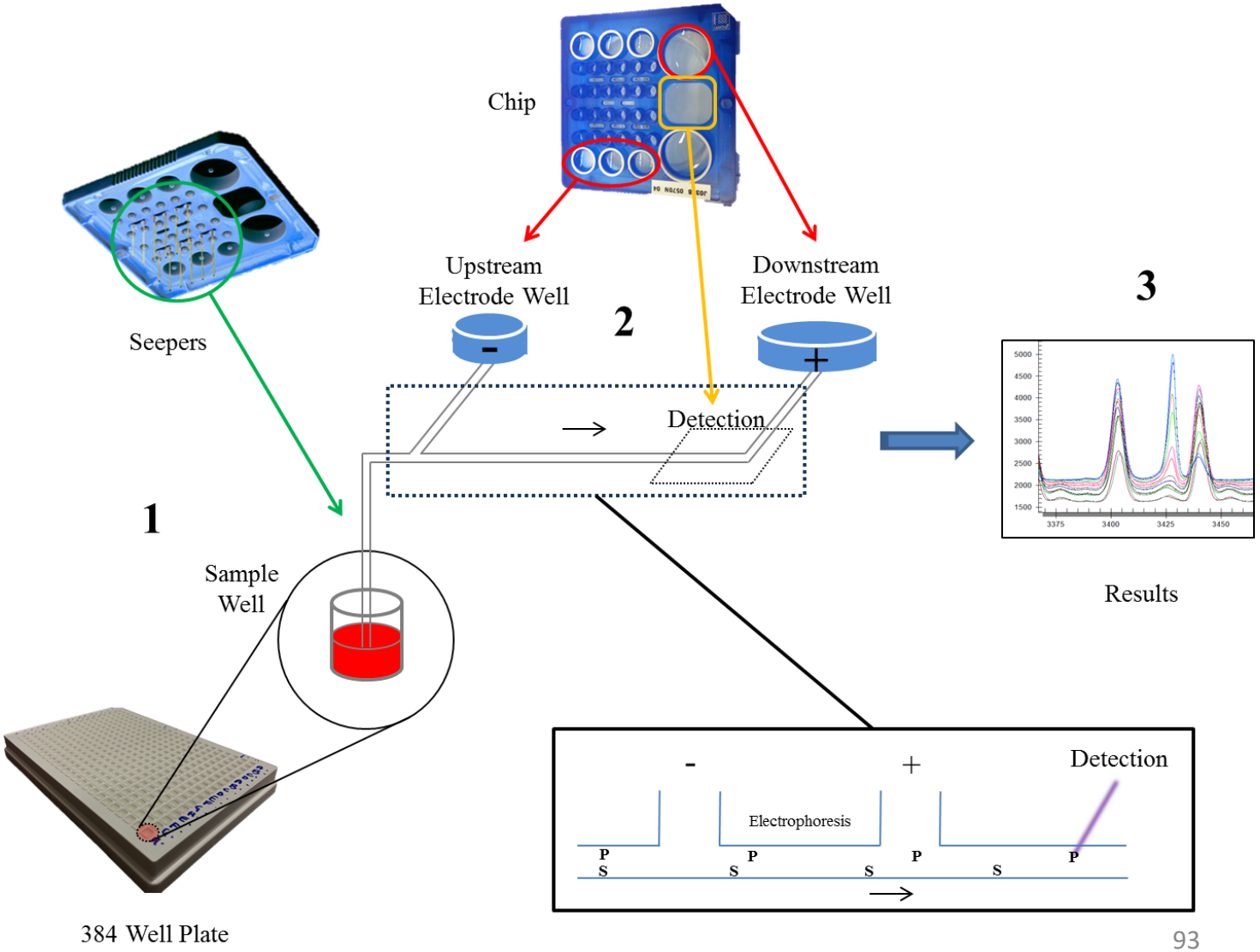
✓ This model should provide a useful tool for the a priori prediction of activity of compounds yet to be synthesized in order to improve their selectivity profiles

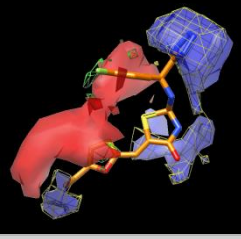
RESEARCH PERIOD ABROAD
(08/2013-10/2013)

Biological evaluation of new Histone Deacetylases Inibitors (HDACIs) against HDAC 3 and HDAC 6





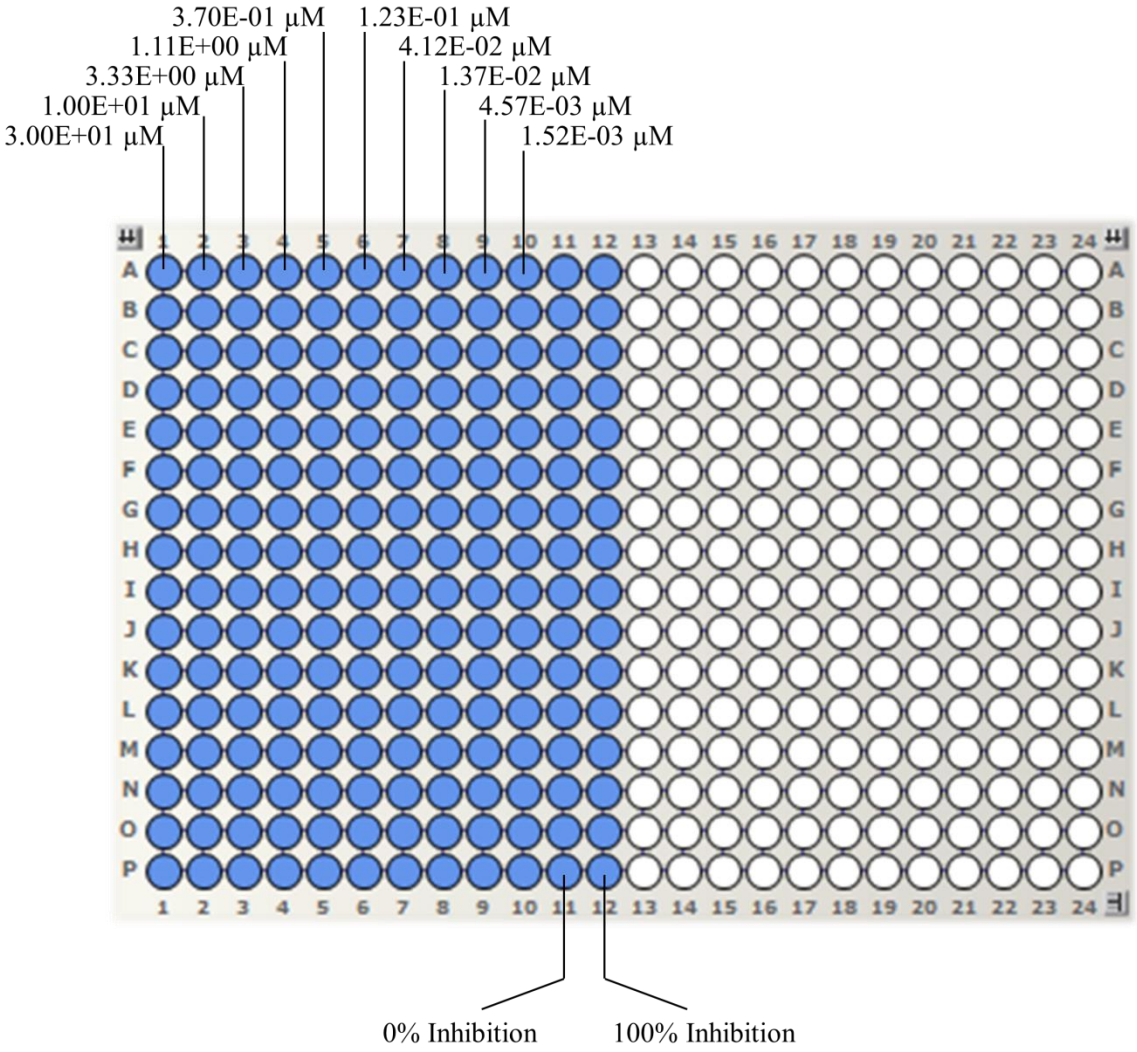




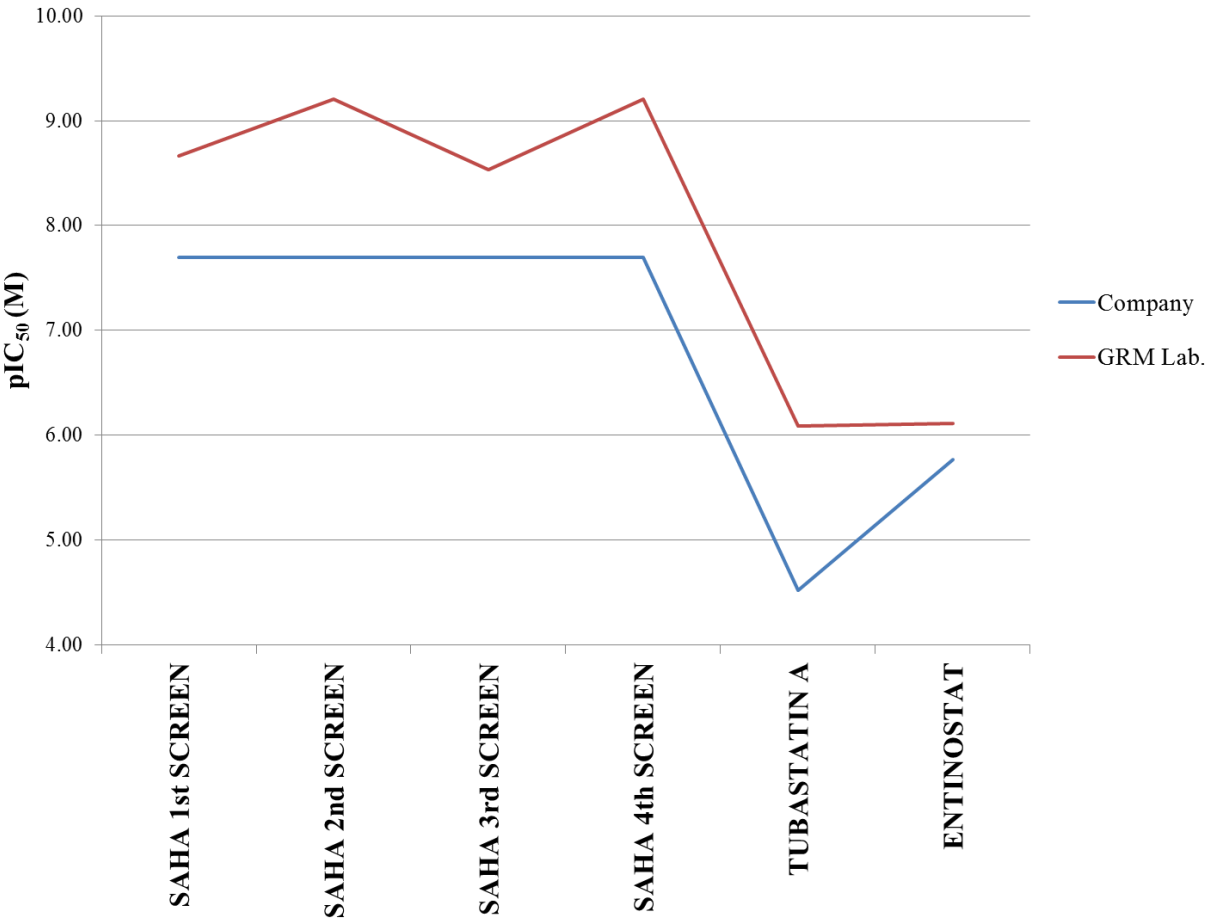
List of screened compounds

#	Cpd	Stock Solution mM in 100% DMSO	#	Cpd	Stock Solution mM in 100% DMSO
1	MC1716	10	16	MC2776	10
2	MC1723	10	17	MC2780	10
3	MC1739	10	18	MC2984	10
4	MC1742	10	19	MC3004	20
5	MC1746	10	20	MC3031	10
6	MC1862	10	21	MC3050	10
7	MC2122	20	22	MC3079	10
8	MC2126	10	23	SD-L-148	20
9	MC2129	10	24	SD-L-256	10
10	MC2195	20	25	SDM141	20
11	MC2427	10	26	SDM146	20
12	MC2625	10	27	ENTINOSTAT[a]	50
13	MC2664	10	28	SAHA[a]	100
14	MC2726	10	29	TUBASTATIN A[a]	100
15	MC2727	10			

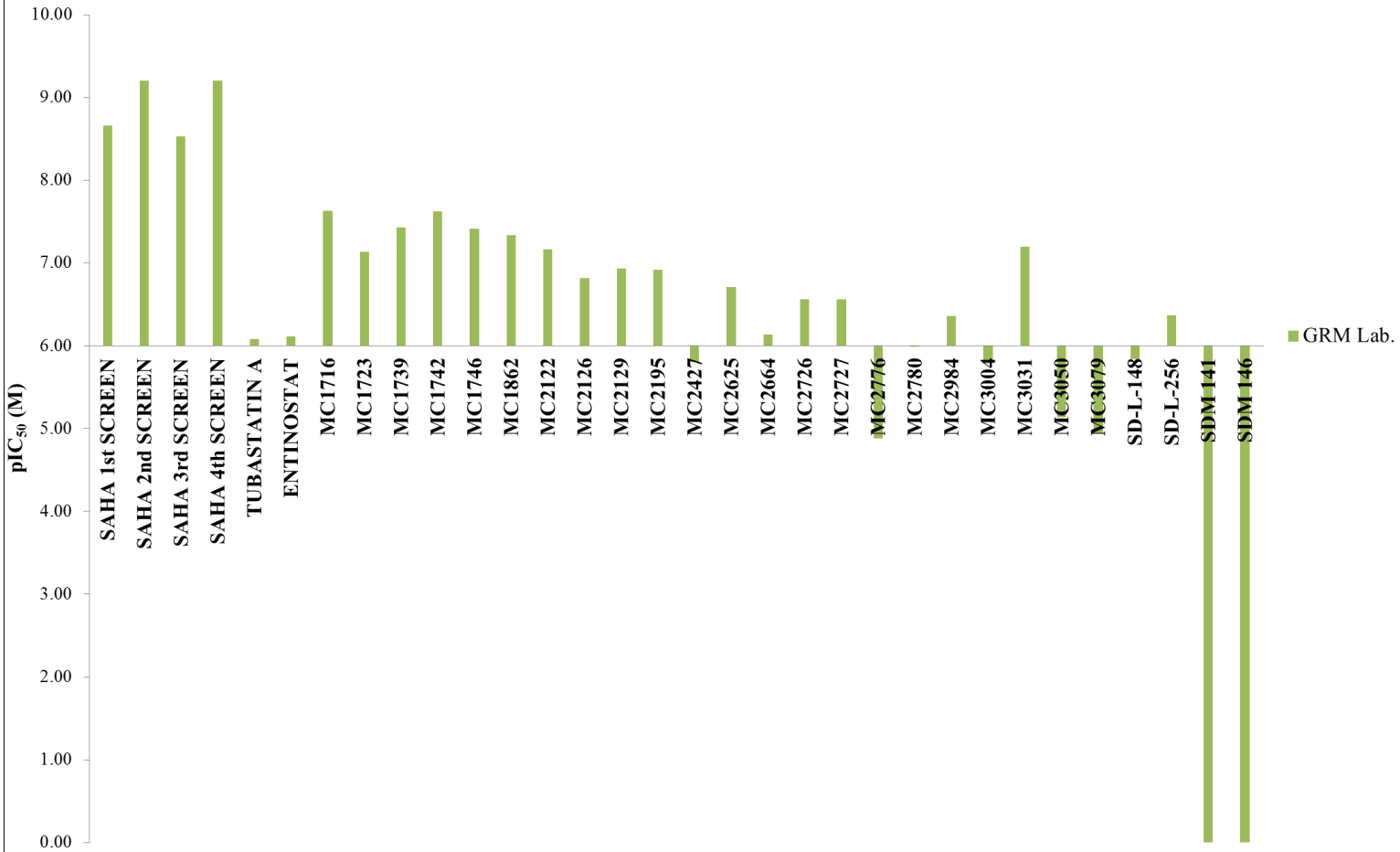
[a]Stock compounds: from Selleckchem® Tubes



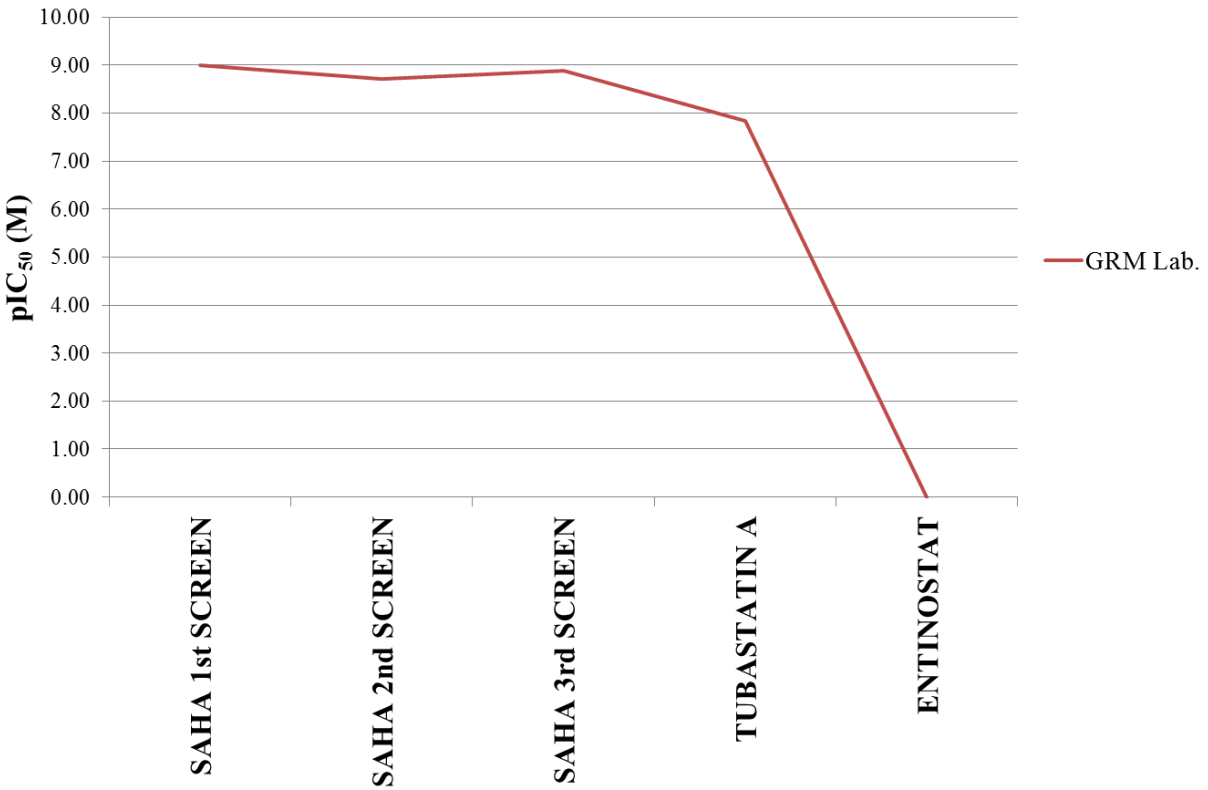
HDAC 3: standard compounds' inhibitory profiles



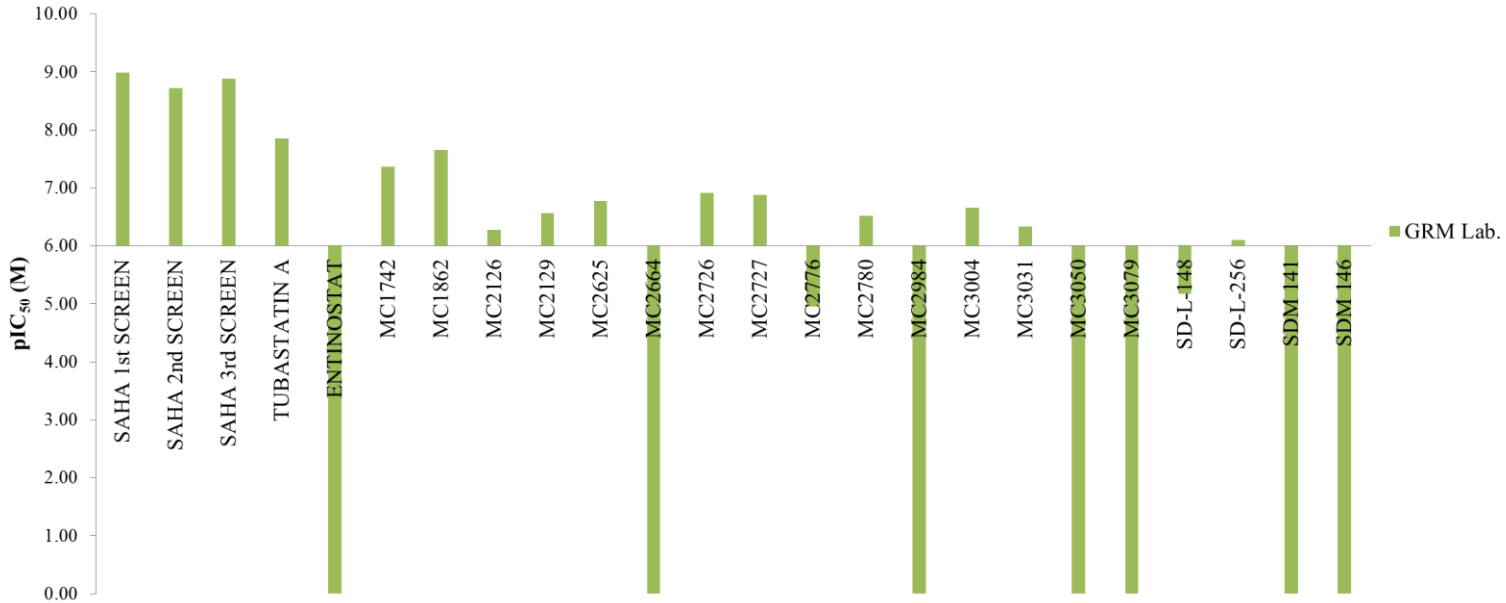
HDAC 3: compounds' inhibitory profiles



HDAC 6: standard compounds' inhibitory profiles



HDAC 6: compounds' inhibitory profiles



➢ HDAC biological investigation

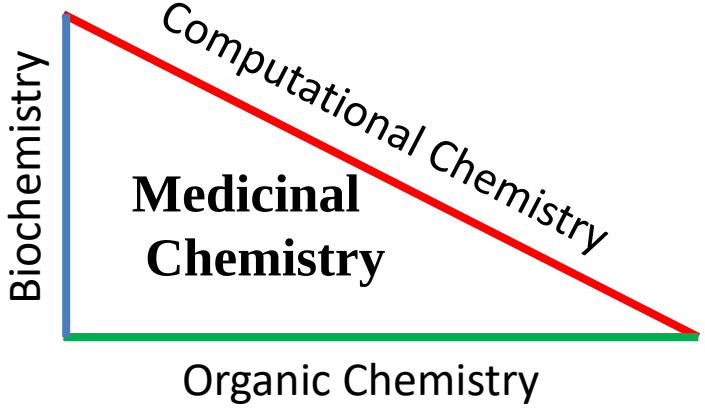
✓ Optimized assay protocol developed

✓ A huge assay investigation is just started…

✓ …over 400 new compounds will be evaluate over the 11 HDACs isoforms

➢ CONCLUSIONS

✓→



✓→ 9 Months Abroad

✓→ 8 Journal Publications + 1 Abstract Published
1 Oral Presentation
14 Scientific Abstract/Poster

✓→ 1 WO Patent (COMBINEr)

Sapienza Università di Roma

Prof. Rino Ragno
Dr. Adele Pirolli
Dr. Alberto De Petris
Dr. Ira Musmuca
Alexandros Patsilinakos
All the RCMDers!

Prof. Antonello Mai
Dr. Sergio Valente
Dr. Dante Rotili

Prof. Enzo Agostinelli
Prof. Vincenzo Tumiatti
Prof. Mariangela Biava
Prof. Silvano Tortorella

Washington University in St. Louis
School of Medicine

Prof. Garland R. Marshall
Dr. Chris Ho
Dr. Peter Harris
Dr. Jayne Marasa
Dr. Nadarapu Damodara Reddy
Richard B. Wickersham III
Sukrith Singh

Université de Lorraine

Prof. Gilbert Kirsch
Prof. Stéphanie Hesse
Dr. Charlène Gadais
Dr. Zhanjie Xu
Dr. Enrico Perspicace

Thanks for Your attention

SAPIENZA
UNIVERSITÀ DI ROMA

SAPIENZA
UNIVERSITÀ DI ROMA