PLATO

Polypharmacology pLATform for predictiOn

Table of Contents

Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it

Preface

What's PLATO?

Standing for Polypharmacology pLATform for predictiOn, PLATO is a ligand-based polypharmacology predictive platform, designed for finding new putative protein drug targets and quantifying bioactivity affinity value. PLATO employs a pool including 611710 druglike ligands provided with experimental bioactivity records retrieved from ChEMBL (release 28, July 2021) according to transparent filtering rules and implements two multi-fingerprint similarity-based predictive algorithms.

Where is PLATO?

PLATO is available at the following address:

http://plato.uniba.it

Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it

Behind the scene

Data sources for PLATO

PLATO employs data available from ChEMBL, a manually curated database of bioactive molecules with drug-like properties.
Two different screening options are given on the basis of two predictive algorithms. The first aims at searching for putative drug targets based on molecular similarity. The second allows making quantitative predictions of bioactivity based on a statistical approach.

Both algorithms share a set of rules to transfer high quality data from ChEMBL to PLATO:

- only records with established biological relationships;
- only small molecules;
- only protein targets;
- only IC_{50}, EC_{50}, and K_i bioactivity data. K_d values are also considered for the drug target prediction algorithm;
- assay confidence score >5
- only relation "=" is considered; specifically records annotated with relations "<=" and ">=" are not considered.
- whenever possible, ChEMBL reports the activity also in standard units 'nM', records reporting concentration in other units are discarded.

ChEMBL coverage

The drug target prediction algorithm is built on a set of 630603 ChEMBL ligands for 5898 targets whereas for quantitative bioactivity profiling is estimated basis on a total of 623980 ChEMBL ligands for 5705 targets.

PLATO algorithms in the literature

Ciriaco, F.; Gambacorta, N.; Alberga, D.; Nicolotti, O.
Quantitative Polypharmacology Profiling Based on a Multifingerprint Similarity Predictive Approach.
(2021) *Journal of Chemical Information and Modeling*, 61(10), pp. 4868–4876. DOI: 10.1021/acs.jcim.1c00498.

Montaruli M., Alberga D., Ciriaco, F., Trisciuzzi D., Tondo A. R., Mangiatordi G. F., Nicolotti O.
Accelerating drug discovery by early protein drug target prediction based on multi-fingerprint similarity search
(2019) *Molecules*, 24(12), 2233
ISSN 1420-3049; DOI:10.3390/molecules24122233

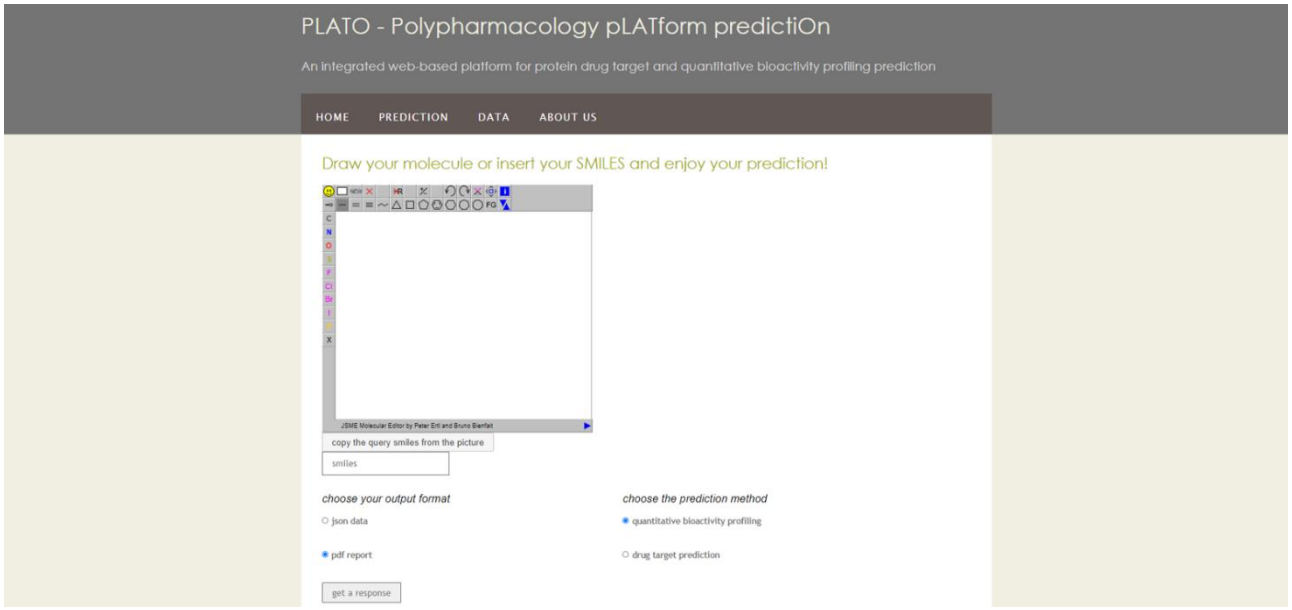
Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it

Alberga D., Trisciuzzi D., Montaruli M., Leonetti, F., Mangiatordi G. F., Nicolotti O.
A new approach for drug target and bioactivity prediction: the Multi-fingerprint Similarity Search aLgorithm (MuSSeL)
(2019) *Journal of Chemical Information and Modeling,* 59(1), pp. 586-596
ISSN: 15499596; DOI: 10.1021/acs.jcim.8b00698

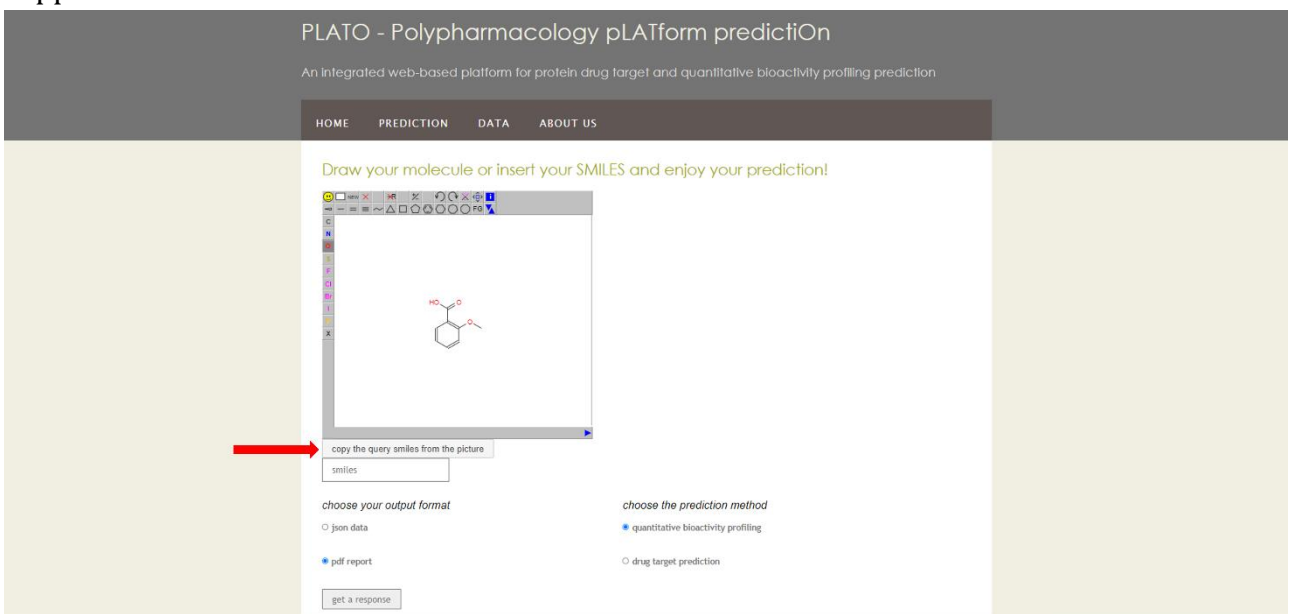
Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it

Getting started with PLATO

Enter the PREDICTION panel.

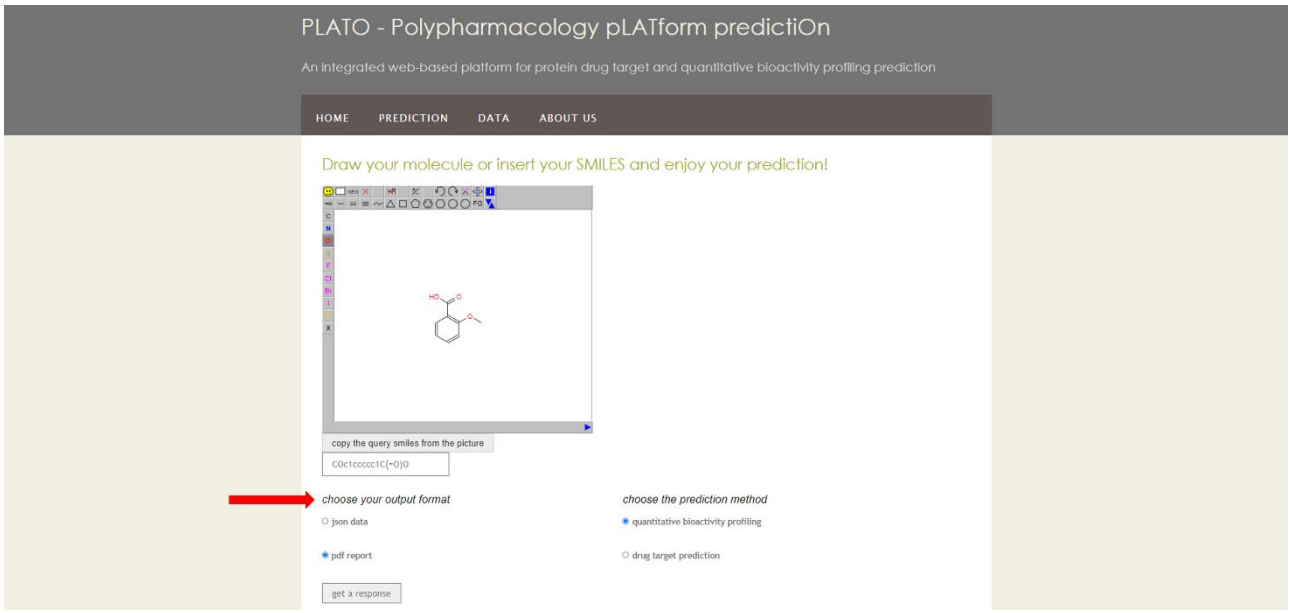


Sketch the query and press "*copy the query smiles from the picture*" to translate the 2D structure in SMILES format.
The user can also paste a SMILES notation into the sketch box; MOL and SDF formats are also supported.



The user can choose which algorithm to use by flagging *"quantitative bioactivity profiling"* or *"drug target prediction"*.
Furthermore, the user can choose the output format as a json file, by flagging *" json data"*.

Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it



Click on button "*get a response*" and enjoy your prediction!

Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it

Read the output and download relevant protein drug target and bioactivity data

- ### Drug target prediction

Based on the algorithm, a downloadable PDF output is returned. In the first page, the name of the algorithm, the query molecule chemical structure and smile notation are reported.
A variable length table is thus generated, containing information about the first 30 predicted targets (default setting).

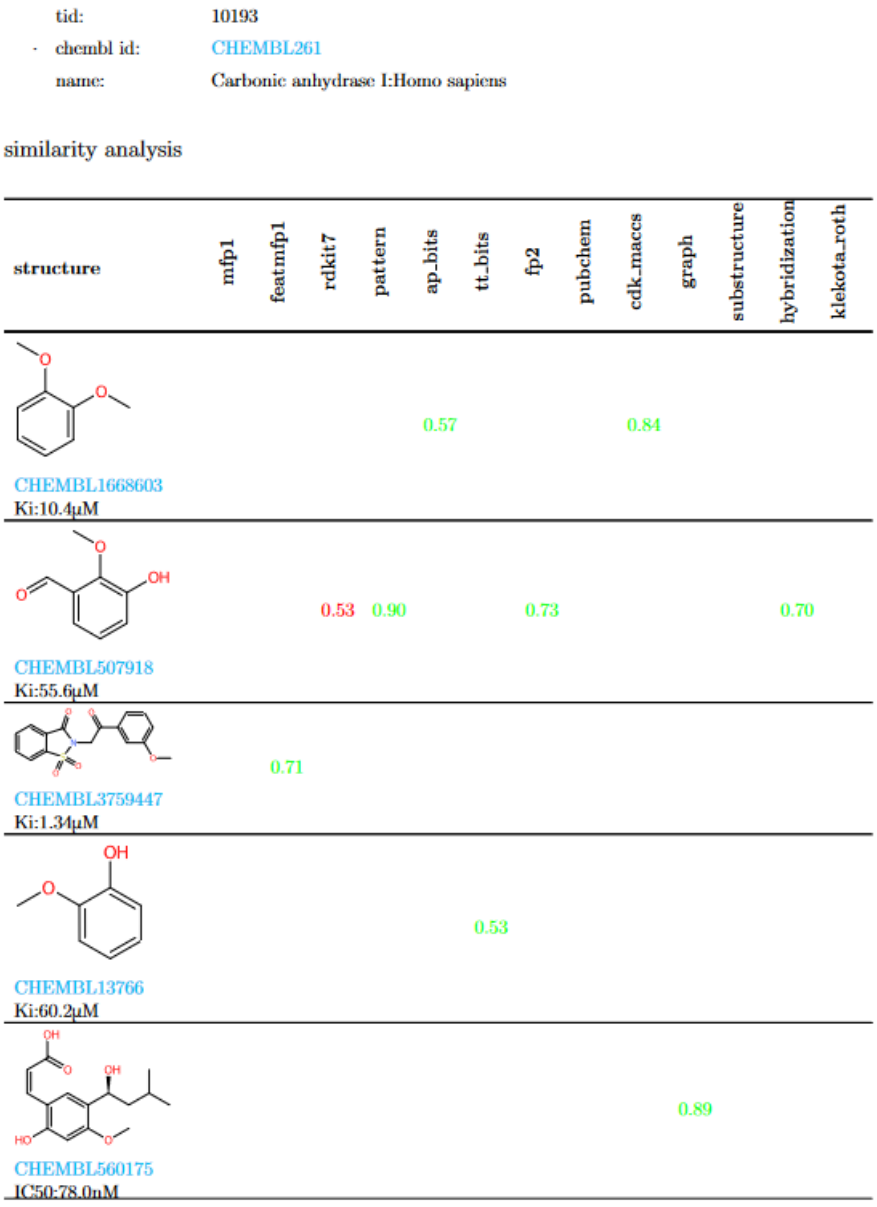
The report of the quantitative bioactivity profiling algorithm contains:

- the 'Target name', from which the user can be forwarded to a page with the prediction details;
- the score, on a scale 0-13 of the protein drug target ranking.
- the 'reliable' column expressing the probability to detect a target with a degree of accuracy according to a similarity threshold.

Target	score	reliable
Carbonic anhydrase I:Homo sapiens	7.40	yes
Carbonic anhydrase IX:Homo sapiens	7.38	yes
Carbonic anhydrase II:Homo sapiens	7.38	yes
Arachidonate 5-lipoxygenase:Rattus norvegicus	7.08	yes
Carbonic anhydrase XII:Homo sapiens	6.73	yes
Serine/threonine-protein kinase/endoribonuclease IRE1:Homo sapiens	6.03	yes
Carbonic anhydrase XIV:Homo sapiens	5.99	no
Carbonic anhydrase VII:Homo sapiens	5.99	no
Tyrosinase:Agaricus bisporus	5.80	no
Cyclooxygenase-1:Homo sapiens	5.75	no
Cyclooxygenase-2:Ovis aries	5.59	no
Cyclooxygenase-1:Ovis aries	5.56	no
Induced myeloid leukemia cell differentiation protein Mcl-1:Homo sapiens	5.43	no
P-glycoprotein 1:Homo sapiens	5.36	no
Acyl coenzyme A:cholesterol acyltransferase:Homo sapiens	5.30	no
Carboxylesterase 2:Homo sapiens	5.29	no
Acyl coenzyme A:cholesterol acyltransferase:Oryctolagus cuniculus	5.26	no
Monoamine oxidase B:Homo sapiens	5.01	no
Glycogen synthase kinase-3 beta:Homo sapiens	4.91	no
Cyclooxygenase-2:Homo sapiens	4.86	no

Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it

Clicking on each target (e.g., Carbonic anhydrase I: Homo Sapiens), a 'similarity analysis' table is obtained reporting the compounds most similar to the query along with the best activity experimental values reported in ChEMBL.
Please note that for each target no more than 13 most similar compounds can be reported.



tid: 10193
chembl id: CHEMBL261
name: Carbonic anhydrase I:Homo sapiens

similarity analysis

structure	mfp1	featmfp1	rdkit7	pattern	ap_bits	tt_bits	fp2	pubchem	cdk_maccs	graph	substructure	hybridization	klekota_roth
CHEMBL1668603 Ki:10.4μM					0.57				0.84				
CHEMBL507918 Ki:55.6μM			0.53	0.90			0.73					0.70	
CHEMBL3759447 Ki:1.34μM		0.71											
CHEMBL13766 Ki:60.2μM						0.53							
CHEMBL560175 IC50:78.0nM										0.89			

The 'similarity analysis' table is organized as follows:

- the Target ChEMBL id directly forwards to the relevant query on the ChEMBL site (i.e., Target Report Card);

Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it

- the 'Structure' column with the ChEMBL identifiers, directly linking to the corresponding ChEMBL page (i.e., Compound Report Card), where the user can recover further information;
- the other 13 columns provide the Tanimoto value of the fingerprint for the ligands for which this attains a maximum value. A green/red flag indicates over/under-threshold similarity values.

- **Quantitative bioactivity profiling**

The quantitative bioactivity profiling report contains a 'summary table' consisting of:

- 'Target name', from which the user can be forwarded to a page with the prediction details;
- Predicted activity values expressed as IC_{50}, K_i or EC_{50};
- σ_p, the variance for the best predicted activity type.

Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it

Target	IC50	Ki	EC50	σ_P
Carbonic anhydrase II:Homo sapiens	78.1nM	56.6nM		1.9: Ki
Carbonic anhydrase I:Homo sapiens	444nM	442nM		1.9: Ki
Carbonic anhydrase IX:Homo sapiens	493nM	73.2nM		2: Ki
Carbonic anhydrase XII:Homo sapiens	1.21µM	41.9nM		2: Ki
Carbonic anhydrase VII:Homo sapiens	5.68µM	94.3nM		2: Ki
Carbonic anhydrase XIV:Homo sapiens	1.02µM	180nM		2: Ki
Cathepsin B:Capra hircus	1.41µM	5.25µM		2: Ki
Plasmepsin 2:Plasmodium falciparum	323nM	48.0nM		2: IC50
Cathepsin D:Homo sapiens	1.48µM	251nM		2: IC50
Carbonic anhydrase IV:Homo sapiens	211nM	616nM		2.1: Ki
Acetylcholinesterase:Homo sapiens	624nM	165nM	581nM	2.1: IC50
Acyl coenzyme A:cholesterol acyltransferase:Homo sapiens	545nM	343nM		2.2: Ki
Serine/threonine-protein kinase/endoribonuclease IRE1:Homo sapiens	438nM		47.4µM	2.2: IC50
Carbonic anhydrase 3:Bos taurus		8.72µM		2.3: Ki
Histone deacetylase 8:Homo sapiens	895nM	134nM		2.4: Ki
Cytochrome P450 1A2:Homo sapiens	2.79µM	739nM	49.6µM	2.4: IC50
Histone deacetylase 4:Homo sapiens	493nM	412nM	2.33µM	2.4: Ki
Carboxylesterase 2:Homo sapiens	1.90µM	193nM		2.4: Ki
Acyl coenzyme A:cholesterol acyltransferase:Oryctolagus cuniculus	14.1nM	198nM		2.4: Ki
Poly [ADP-ribose] polymerase-1:Homo sapiens	125nM	23.1nM	65.7nM	2.4: IC50
Cytochrome P450 2A5:Mus musculus	7.65µM			2.4: IC50
Cytochrome P450 2A6:Homo sapiens	4.30µM	880nM		2.4: IC50
Arachidonate 5-lipoxygenase:Rattus norvegicus	1.38µM	800nM		2.5: IC50
Cyclooxygenase-1:Homo sapiens	5.58µM	31.4µM		2.5: IC50
Receptor protein-tyrosine kinase erbB-2:Homo sapiens	291nM	222nM	113nM	2.5: IC50
Interleukin-1 receptor-associated kinase 4:Homo sapiens	40.1nM	719nM		2.5: IC50
Leukocyte elastase:Homo sapiens	89.2nM	44.3nM	2.67nM	2.5: Ki
Serine/threonine-protein kinase PLK1:Homo sapiens	254nM	538nM	5.51µM	2.5: IC50
Transient receptor potential cation channel subfamily A member 1:Rattus norvegicus	1.95µM		3.11µM	2.5: EC50
Quinone reductase 2:Homo sapiens	902nM	204nM		2.5: IC50

The 'similarity analysis' table reports for each target the compounds most similar to the query along with the mean experimental activity values (IC_{50}, K_i or EC_{50}) reported in ChEMBL.
It is organized as follows:

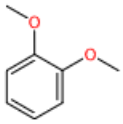
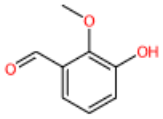
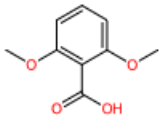
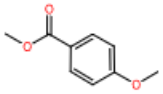
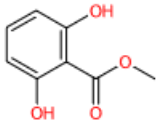
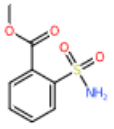
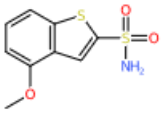
- the Target ChEMBL id can be clicked to retrieve information about the target from the ChEMBL server (i.e., Target Report Card);
- the 'Structure' column with the ChEMBL identifiers link to the corresponding ChEMBL page, where the user can recover further information (i.e., Compound Report Card);.

Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it

- the τ value expressing a value of biological activity precision. The higher τ value, the higher is the molecular similarity in respect to your query.

tid: 15
chembl id: CHEMBL205
name: Carbonic anhydrase II:Homo sapiens

similarity analysis

structure	chembl id	τ	IC50	Ki	EC50
global result		0.053: Ki	78.1nM	56.6nM	
	CHEMBL1668603	0.031		14.9nM	
	CHEMBL507918	0.024		22.1nM	
	CHEMBL488609	0.024		20.3nM	
	CHEMBL1762668	0.018		34.7nM	
	CHEMBL2333586	0.012		17.2nM	
	CHEMBL148680	0.012	130nM		
	CHEMBL315921	0.0077	2.94nM		

Alternatively, a json output can be downloaded. This is mainly for inspection by programs but contains many more details both about the prediction mechanics and about the compound properties.

Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it

Word of caution

The current release of PLATO has been just updated (December 2021). The authors have made much effort to provide a finished product. Please report any omission, interface bug or any result that strikes you as odd or wrong.
Your feedback is important to us!

Orazio Nicolotti
Dipartimento di Farmacia-Scienze del Farmaco
Università di Bari 'Aldo Moro' (Italy)
orazio.nicolotti@uniba.it