



Open PHACTS – The use of open data for *in silico* models

Gerhard F. Ecker Dept. of Pharmaceutical Chemistry, Univ Vienna <u>Gerhard.f.ecker@univie.ac.at</u>; www.openphacts.org







Approaching complex research questions needs integration of data sources

Open Pharmacological Space The Need



Number	sum	Nr of 1	Question
15	12	9	All oxidoreductase inhibitors active <100nM in both human and mouse
18	14	8	Given compound X, what is its predicted secondary pharmacology? What are the on and off,target safety concerns for a compound? What is the evidence and how reliable is that evidence (journal impact factor, KOL) for findings associated with a compound?
24	13	8	Given a target find me all actives against that target. Find/predict polypharmacology of actives. Determine ADMET profile of actives.
32	13	8	For a given interaction profile, give me compounds similar to it.
37	13	8	The current Factor Xa lead series is characterised by substructure X. Retrieve all bioactivity data in serine protease assays for molecules that contain substructure X.
38	13	8	Retrieve all experimental and clinical data for a given list of compounds defined by their chemical structure (with options to match stereochemistry or not).
41	13	8	A project is considering Protein Kinase C Alpha (PRKCA) as a target. What are all the compounds known to modulate the target directly? What are the compounds that may modulate the target directly? i.e. return all cmpds active in assays where the resolution is at least at the level of the target family (i.e. PKC) both from structured assay databases and the literature.
44	13	8	Give me all active compounds on a given target with the relevant assay data
46	13	8	Give me the compound(s) which hit most specifically the multiple targets in a given pathway (disease)
59	14	8	Identify all known protein-protein interaction inhibitors

Azzaoui et al., Drug Discov Today 2013, 843-52



The User



The Explorer is a user-friendly, full featured interface that allows scientists to explore and interrogate integrated biological and chemical data

The API allows searches along the concepts of

compound – target – pathway - disease









Explorer Search... Sildenafil Pharmacology Data View in ChemBioNavigator Sidenahi (in citizate form), sold under the names Vagra, Bweitio and under vortous other names, is a drug used to treat male exettle dydnuction (impotence) and guinomary activatial hypetencion (PAH), developed by the pharmaceutical company Pfizer. Its primary competitors on the market are tadalafi (Caalis), and verdenahi (Levira), (Wupetai) Hepatic # H-Bond Receptors: ChemSpider ID: 5023 Molecular Formula: C22H30N6O4S # H-Bond Donors SMILES: O=S(=O)(N1CCN(C)CC1)c4cc(C)2=N(C(=O)c3c(N/2)c(nn3C)CCC)c(OCC)cc4Standard InChl: $\label{eq:linear} InchI=15/C22H30N6O45/c1-5-7-17-19-20(27(4)25-17)22(29)24-21(23-19)16\\ 14-15(8-9-18(16)32-6-2)33(30,31)28-12-10-26(3)11-13-28/h8-9,14H,5-7,10-13H2,1-4H3,(H,23,24,29)\\ 13H2,1-4H3,(H,23,24,29)\\ \end{tabular}$ Mol Weight: BNRNXUUZRGQAQC-UHFFFAOYSA-N 474,576 Standard InChiKey: Affected Organi Humans and other mammals MAN Emphas 474,576 Indication: For the treatment of erectile dysfunction 189-190 oC Melting Point: Polar Surface Area 117.51 # Rotatable Bonds:

2.2

7

1

7







STANDARD_TYPE	UNIT_COU	NT			
AC50	7	STAI	NDARD_TYPE	STANDARD_UNITS	COUNT (*)
Activity	421	IC5	C	nM	829448
EC50	39	IC5	C	ug.mL-1	41000
	40	IC5	C		38521
Ki	42	IC5	C	ug/ml	2038
Log IC50	4	IC5	C	ug ml-1	509
Log Ki	7	IC5	0	mg kg-1	295
Potency	11	IC5	0	molar ratio	178
$\log TC50$	0	IC5	C	ug	117
109 1000	0	IC5	C	00	113
		IC5	C	uM well-1	52
5000 turno		IC5	C	p.p.m.	51
>ooo type:	5	IC5	C	ppm	36
		IC5	C	uM-1	25
		IC5	C	nM kg-1	25
		IC5	C	milliequivalent	22
		IC5	C	kJ m−2	20

Implemented using the Quantities, Dimension, Units, Types Ontology (http://www.qudt.org/)

~ 100 units





This is the first public release of the Open PHACTS Explorer, and we look forward to and value your feedback and comments .











(+)

TSV Downloads







	Compound by name 🗵	Pharmacology by Co	npound name 🗵											
on 🦳	Н	nt: Type in compou	nd name. E.g. "Aspirin"	and select a result										
und	Compound name: P	opafenone				Search								
npound by name	Filter Dr.													
npound by structure	Pileo	venance: O O	ι 🔘 Οπ											
	Pharmacology by Compo	nd name search resu	lts - Total Records: 122											
Pharmacology	Repare tsv file													
rmacology by Enzyme family	Structure	Compound Name	Target Names	Target Organisms	Assay Organism	Assay Description	Activity Type R	elation	Value	Units	SMILES	InChi	InChi Key	Pub
Pharmacology by Compound Pharmacology by Target	1 milet	Propafenone	P-glycoprotein 1	Homo sapiens		Compound was tested for inhibition of daunomycin efflux in the resistant human T-lymphoblast cell line CEM vcr1000.	EC50	=	320	nM	0=C(c1ccccc10	InChI=15/C21H	JWHAUXFOSRP	<u>97</u>
	2 And the second second	Propafenone	L5178Y (Lymphoma cells)	Mus musculus	Mus musculus	Compound was tested for inhibition of rhodamine 123 efflux in mdr 1 transfectant L5178Y VMDRI C.06 mouse lymphoma cells; nd=Not determined					0=C(c1ccccc10	InChI=15/C21H	JWHAUXFOSRP	<u>97</u> (
	3 marile	Propafenone				Calculated membrane partition coefficient (Kmemb)	Log Kmemb	-	0.7		0=C(c1ccccc10	InChI=15/C21H	JWHAUXFOSRP	<u>150</u>
	4 marilie	Propafenone				Volume of distribution in man (IV dose)	Vdss	-	3,6	L.kg-1	0=C(c1ccccc10	InChI=1S/C21H	JWHAUXFOSRP	120
eedback +	5	Propafenone				Unbound fraction (plasma)	Fraction fr		0.05		0=C(c1ccccc10	InChI=1S/C21H	JWHAUXFOSRP	120





~	Compound by na	ame 🗵 🏻 Pharm	acology by Compou	nd name 🗵 Compound Structure Search 🗵 Target	by name 🗵			
Navigation -		Hint: Sta	art typing in prote	in name and species. E.g. "Adenosine receptor A	2a (Homo sapiens)"			
🔁 Compound	Target name:	rget name: Multidrug resistance protein 1A (Mus musculus) Search						
E Compound by name	Target Data							
Compound by structure								
The second secon								
E larget by name	1000	AND						
Pharmacology	A A A		M. detalu	nun unsistemen nustain 2	(Mus managed			
Pharmacology by Enzyme Tarmiy	NRA.	Lagar	Multiar	rug resistance protein 3	(Mus muscul	us)		
Pharmacology by Compound								
	Caral	and a	Pharmacology Da	ata				
	- and the	C 4 2 9						
		a sta	Description:	Multidrug resistance protein 3				
			Synonyms:	ATP-binding cassette sub-family B member 1A, CHEM protein 3, P-glycoprotein 3	BL2573, MDR1A, Multidrug resist	ance		
			Specific Function	: Energy-dependent efflux pump responsible for decre resistant cells.	ased drug accumulation in multi	lrug-		
			Keywords:	Hydrolase, Phosphoprotein, 3D-structure, ATP-bindin proteome, Direct protein sequencing, Glycoprotein, I Reference proteome, Repeat, Transmembrane, Tran	g, Cell membrane, Complete Membrane, Nucleotide-binding, smembrane helix, Transport			
			PDB Entry:	3G61 3G5U 3G60				
	5							
neip and reedback								
API Status +								
TSV Downloads +								





Open PHACTS Ex	plorer															
() ()	Compour	nd by name 🗵 🏾 Pl	harmacology by Compound name	Compound Structure Search	Target by name	Compounds	active against enzyme famil	y × Pharm	acology b	y Target I	Name 🗵					
Navigation -		Hint	:: Type in protein name and s	pecies. E.g. "ADA protein human'	and select a resul	t										
a 🔄 Compound	Protein n	ame: Mult	idrug resistance protein 1 (Homo :	sapiens)		Search										
Compound by name Compound by structure	O Filte	r Prove	nance: 💿 On 💿 Off 🧃													
⊿ 🔂 Target	Pharmaco	ology by Target na	me search results - Total Records	: 2563												
E larget by name	Prepare	e tsv file														
Pharmacology by Enzyme family	Str	ructure	Compound Name	Target Name	Target Organism	Assay Organism	Assay Description	Activity Type	Relation	Value	Units	Mol Weight	SMILES	InChi	InChi Key	
 Pharmacology by Compound Pharmacology by Target 	1	Ø					Compound was tested for									
	c.		4-benzoyl-1-[2-hydroxy- ³⁴) 3-(propan-2-ylamino)propyl]-5- methyl-2-phenyl-1,2-dihydro- ³ H-pyrazol-3-one	Multidrug resistance protein 1 (Homo sapiens)	Homo sapiens		inhibition of daunomycin efflux in the resistant human T-lymphoblast cell line CEM vcr 1000.	EC50	=	83180	nM	393.479	0=C2\C(=C(/N(InChI=1S/C23H	VZPAOCHCLZMT	
	2 H ₈		I, 3-dimethyl- 4-{(E)-phenyl(propylimino)methyl pyrazol-5-ol	Multidrug resistance protein 1 (Homo sapiens)	Homo sapiens		Compound was tested for inhibition of daunomycin efflux in the resistant human T-lymphoblast cell line CEM vcr 1000.	EC50	=	272500	nM	257.331	CCC/N=C(\clcc	InChI=1S/C15H	gjpfqwwtopa	
	3		H, 4-benzoyl-1-[3-(dipropan- 2-ylamino)-2-hydroxypropyl]-2,5- dimethyl-1,2-dihydro- 3H-pyrazol-3-one	Multidrug resistance protein 1 (Homo sapiens)	Homo sapiens		Compound was tested for inhibition of daunomycin efflux in the resistant human T-lymphoblast cell line CEM vcr1000.	EC50	=	24120	nM	373,489	O=C1\C(=C(/N(InChI=1S/C21H	UYZSJNDVHQW	
	4		1-[2+Hydroxy- 3-(propylamino)propyl]-5-methyl- 2-phenyl- 4-(2-thienylcarbonyl)-1,2-dihydro- 3H-pyrazol-3-one	Multidrug resistance protein 1 (Homo sapiens)	Homo sapiens		Compound was tested for inhibition of daunomycin efflux in the resistant human T-lymphoblast cell line CEM vcr1000.	EC50	3	72440	nM	399.507	O=C(C=2C(=O)	InChI=1S/C21H	JCKJIYACBBEC	
	5	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	1-[2-hydroxy-3-(propan- 2-ylamino)propyl]-5-methyl- #, 2-phenyl- 4-(3-ohenylnronanovl)-1.2-	Multidrug resistance protein 1 (Homo sapiens)	Homo sapiens		Compound was tested for inhibition of daunomycin efflux in the resistant human T-lymnhohast cell line CFM	EC50	5	11750	nM	421.532	0=C2\C(=C(/N(InChI=1S/C25H	DNSZEGKDVYSK	
Help and Feedback +		CH,	dihydro-3H-pyrazol-3-one				va 1000.									
API Status +		C														
TSV Downloads +	6	-														

Provenance



Open PHACTS I	Expl	lorer							
	«	Home	Pharmacolog	jy by Ta	rget Name 🛎				
Navigation	•			Hint: -	Гуре in protein name and s	pecies. E.g. "Adenosine rece			
 Compound Target Pharmacology 		Protein name:		Multidrug resistance protein 1 (Homo sapiens) Provenance: On Off					
Pharmacology by Enzyme family Pharmacology by Compound		Pharma	cology by Targe	et name	search results - Total Record	s: 5204			
Pharmacology by Compound Pharmacology by Target	E	Prepare tsv file							
		-	Structure		Compound Name	Target Name			
	2			4-1 3-1 3H 3H	penzoyl-1-[2-hydroxy- propan-2-ylamino)propy[]-5- ethyl-2-phenyl-1,2-dihydro- -pyrazol-3-one	P-glycoprotein 1			
		2		сн, 1,, 4- ру	3-dimethyl- :(E)-phenyl(propylimino)methyl]- razol-5-ol	P-glycoprotein 1			
		3		4-1 2	penzoyl-1-[3-(dipropan- /lamino)-2-hydroxypropy[]-2,5-	P-glycoprotein 1			
TSV Downloads	Ð	•							

Open Pharmacological Space

Open Pharmacological Space Target Taxonomy

TSV Downloads



Enzyme family class: No enzyme class selected - pre	Star	e EC codes						
Filter Provenance: On On Pharmacology by Enzyme Family search results Prepare tay file	Star	tsearch						
Filter Provenance: On On On Pharmacology by Enzyme Family search results Prepare tay file	Off 🕑	Calact an annuma funiti						
Pharmacology by Enzyme Family search results	Off	Colort an annuma family						
Pharmacology by Enzyme Family search results		Salact an annuma familia						
(B) Prepare tsv file		Sector 1 and cars structure	Calact an annun famile					
		Select an enzyme family						
		EC number 🔺	Enzyme family name					
Structure Compound Name	Target Name	Þ 🛄 1,-,-,-	Oxidoreductases	▲ ation	n			
		a 🔂 2	Transferases					
		▷ 🧰 2.1,-	Transferring one-carbon groups					
		▷ 🦲 2.10	Transferring molybdenum- or tungsten-containing groups					
		▷ 🧰 2.2	Transferring aldehyde or ketone residues					
		⊿ 🔄 2.3	Acyltransferases					
		▷ 🛄 2.3.1	Transferring groups other than amino-acyl groups					
		⊿ 🔄 2.3.2	Aminoacyltransferases					
		≘ 2.3.2.1	D-glutamyl transpeptidase,D-glutamyltransferase	E				
		2.3.2.10	UDP-N-acetylmuramoylpentapeptide-lysine N(6)-alanyltr					
		= 2,3,2,11	Alanyiphosphatidyigiycerol synthase					
ler Dianneo		E 2.3.2.12	Peptidyitransterase					
	E 2.3	2.3.2.13	Giutaminyipepude gamma-giutamiyiransierase, Fibrinolig					
contor toxonomy	1	= 2,3,2,14	Glutathione gamma-glutamylcysteinyltransferase Phyto					
isponer taxonomy		= 2.3.2.15	Linid II divoite divoitransferase					
		= 2.3.2.17	N-acetylmuramovi-H-alanyi-D-glutamvi-H-lysvi-(N(6)-glyc					
eptor taxonomy		= 2,3,2,18	N-acetylmuramovi-L-alanyl-D-glutamyl-L-lysyl-(N(6)-trigl					
		= 2.3.2.2	Gamma-glutamyl transpeptidase.Gamma-glutamyltransf					
JE taxonomy		= 2.3.2.3	Lysyltransferase					
		= 2.3.2.4	Gamma-glutamylcyclotransferase					
		= 2.3.2.5	Glutaminyl cyclase, Glutaminyl-peptide cyclotransferase,					
		2.3.2.6	Leucyl-tRNAprotein transferase,L/F transferase,Leuc					
		2.3.2.7	Aspartyltransferase					
		2.3.2.8	Arginyltransferase,Arginyl-tRNAprotein transferase					
	ner planned nsporter taxonomy ceptor taxonomy ME taxonomy	ner planned nsporter taxonomy ceptor taxonomy ME taxonomy	Image: Spectrum	 Participant and the second seco	Image: state in the state			



-



File Edit Select Render Protein Compute Window Help

Moe 2012.10

			_					Open
Database Viewer : d:/1304 - april 2013/example2.mdb	- 🗆 💌 X 🚥							LigX
File Edit Display Compute OpenPHACTS Window Help SVI DBV MOE	Cancel							Constrain •
Pharmacology By Target								Close
Pharmacology By Compound								Center
Pharmacology By <u>E</u> nzyme Family								SiteView
Pharmacology By ChEBI Class								Hydrogens
								Show
								Ligand
Input CHEMBL-target ID								Surface
Please input the CHEMBL ID	of your target							Measure •
ОК	Cancel							Builder
								Sketch
								Minimize
	Database Viewer : d:/1304 - april	2013/example.mdb						Select •
0 entries. 1 field, 0 selected, all visible.	<u>File Edit Display Compute Oper</u>	PHACTS Window Help				SVL DBV MOE	Cancel	Extend
	smiles	http://data.kasabi.com/datase	pmid	full_mwt	http://www.conceptwiki.org/	prefLabel	prefLabe	Delete
		http://data.kasabi.com/da	9767638	393.4790	http://www.conceptwiki.	Multidrug resistance protein 1 (Homo sapie	Multidrug	
	- · ·							
	2	http://data.kasabi.com/da	9767 <mark>6</mark> 38	257.3310	http://www.conceptwiki.	Multidrug resistance protein 1 (Homo sapie	Multidrug	
	3 other	http://data.kasabi.com/da	9767638	373 .4 890	http://www.conceptwiki.	Multidrug resistance protein 1 (Homo sapie	Multidrug	
	4 000	http://data.kasabi.com/da	9767638	399.5070	http://www.conceptwiki.	Multidrug resistance protein 1 (Homo sapie	Multidrug	
	····							
	Â							
	5 4.00	http://data.kasabi.com/da	9767638	421.5320	http://www.conceptwiki.	Multidrug resistance protein 1 (Homo sapie	Multidrug	
	~~~~							
	6 Sam	http://data_kasabi_com/da	9767638	393 4790	http://www.concentwiki	Multidrug resistance protein 1 (Homo samie	Multidrug	
	° ~~*		5767656	55511150	incepti)/ initiconceptilitie		hareranag	
				101 5300				
	್ ನ್	nccp://uaca.kasabi.Com/da	9/0/038	421.0020	nccp.//www.conceptW1K1.	Indicional resistance process i (Homo Sable	There is a second secon	
	4	' III '						68 📼
	2563 entries, 0 selected, all visible.	19 fields, 0 selected, all visible.						R S L
								0

MOE

# Open PHACTS Example applications



## **Advanced analytics**

ChemBioNavigator	Navigating at the interface of chemical and biological data with sorting and plotting options
TargetDossier	Interconnecting Open PHACTS with multiple target centric services. Exploring target similarity using diverse criteria
PharmaTrek	Interactive Polypharmacology space of experimental annotations
UTOPIA	Semantic enrichment of scientific PDFs
Pred	ictions
GARFIELD	Prediction of target pharmacology based on the Similar Ensemble Approach
eTOX connector	Automatic extraction of data for building predictive toxicology models in eTOX project



**PharmaTrek** 



### PHARMATREK

TARGETS 🥄	Interaction Map
p38 alpha homo You have 1 targets selected Mitogen-activated protein kinase Mitogen-activated protein kinase 14	
14 (Homo sapiens)     (Homo sapiens)       Amino Acid, Peptide, or Protein     Amino Acid, Peptide, or Protein       Image: Im	
alpha thalassemia/mental retardation syndrome X-linked homolog (human) protein, mouse Amino Acid, Peptide, or Protein	
	Image: Second
📰 🝸 🔀 🗆 connect	
LIGANDS 2(1H)-quinazolinone, 5-(2-chloro-4-fluorophenyl)-1-(2,6- dichlorophenyl)-3,4-dihydro-7-(4-piperidinyl)-	
2(1H)-quinazolinone, 5-(2-chloro-4-fluorophenyl)-1-(2,6- dichlorophenyl)-3,4-dihydro-7-(4-piperidinyloxy)-	$\begin{array}{cccccccccccccccccccccccccccccccccccc$
Image:	

### Open PHACTS Open Pharmacological Space ChemBioNavigator









# COLLECTOR, AN APPLICATION FOR GATHERING BIOACTIVITY DATA FROM THE OPS PLATFORM

# Manuel Pastor











RESEARCH PROGRAMME ON BIOMEDICAL INFORMATICS





### Collector: architecture overview



### Pharmacoinformatics Research Group Department of Pharmaceutical Chemistry



- Transmembrane receptor
- Tetramer
- Non-selective cation channel
- Peripheral nervous system
- Activation → influx of cations → depolarization of membrane → pain perception
- Prolonged activation desensitization



TRPV1



### Pharmacoinformatics Research Group





Department of Pharmaceutical Chemistry

Firefox  Louis Armstrong New	w Orleans Interna × 🚹	i Google Kalender	< 🔥 Register for Open PHACTS Exp	lorer 🛛 🗙 Op	en PHACTS	× +		- 0	X				
explorer.openphacts.org/#!p=Ph	armByTargetNameForm8	&u=http://www.conceptwiki.org/concept/	f1c37c53-63c3-4107-a688-12d4dc082e	:f3	☆ ⊽ C	8 - molecular informatics	; journal 🔎	<b>^</b>	-				
Open PHACTS Ex	cplorer												
×	Pharmacology by Ta	rget Name 🗵											
Navigation 📃		Hint: Type in protein name and sp	ecies. E.g. "ADA protein human	and select a res	sult								
Compound	Protein name:	Transient receptor potential cation cha	nnel subfamily V member 1 (Homo s	apiens)	Search.								
<ul> <li>Target</li> <li>Pharmacology</li> </ul>	O Filter	Provenance: On On Off											
Pharmacology by Enzyme family Pharmacology by Compound	Pharmacology by Tar	acology by Target name search results - Total Records: 2328											
Pharmacology by Target	Prepare tsv file												
	Structure	Compound Name	Target Name	Target Organism	Assay Organism	Assay Description	Activity Type Relation	Valu	e U				
		(3beta)-3-hydroxy-N-(4-hydroxy- 3-methoxybenzyl)urs-12-en- 28-amide	Transient receptor potential cation channel subfamily V member 1 (Homo sapiens)	Homo sapiens		Percent absolute effect obtained with ionomycin, measured by the entry of [Ca2+] into human embryonic kidney HEK293 cells overexpressing hVR1	Response =	14.3	3				
		(3beta)-3-hydroxy-N-(4-hydroxy- 3-methoxybenzyl)urs-12-en- 28-amide	Transient receptor potential cation channel subfamily V member 1 (Homo sapiens)	Homo sapiens	Homo sapiens	[Ca2+] influx into human embryonic kidney HEK293 cells overexpressing human VR1 (NM = non-measurable)							
Help and Feedback	3		Transient recenter patronial estion			Percent absolute effect obtained with ionomycin,							
API Status	-gr-l	(12R)-12-hydroxy-N-(4-hydroxy- 3-methoxybenzyl)octadecanamide	channel subfamily V member 1	Homo sapiens		measured by the entry of [Ca2+] into human	Response =	69.7	7				
TSV Downloads						embryonic kidney HEK293 cells overexpressing hVR 1			-				
			III					_					

Pharmacoinformatics Research Group TRPV1

### hTRPV1 $\rightarrow$ 2328 ligands from Open PHACTS



universität wien

Goldmann D, Mol Inform 2013.

### **Pharmacoinformatics Research Group**









molecular informatics



DOI: 10.1002/minf.201200059

#### Annotating Human P-Glycoprotein Bioassay Data

Barbara Zdrazil,^[a] Marta Pinto,^[a] Poongavanam Vasanthanathan,^[a] Antony J. Williams,^[b] Linda Zander Balderud,^[c] Ola Engkvist,^[c] Christine Chichester,^[d] Anne Hersey,^[e] John P. Overington,^[e] and Gerhard F. Ecker^{#[a]}

### Develop assay ontology Create tools for semiautomatic combination Propose reference compounds



HOME

### The Open PHACTS Foundation

OPF is a not-for-profit membership organisation, supporting the Open PHACTS Discovery Platform:

A sustainable, open, vibrant and interoperable information infrastructure for applied life science research and development.

To reduce the barriers to drug discovery in industry, academia and for small businesses, the Open PHACTS Discovery Platform provides tools and services to interact with multiple integrated and publicly available data sources. To integrate this data, extensive cross-referencing of scientific concepts is needed across all databases.

The Open PHACTS Foundation ensures the sustainability of the Open PHACTS Discovery Platform infrastructure and acts as a hub for relevant scientific research and development.







### **Open PHACTS Project Partners**

### www.openphacts.org

