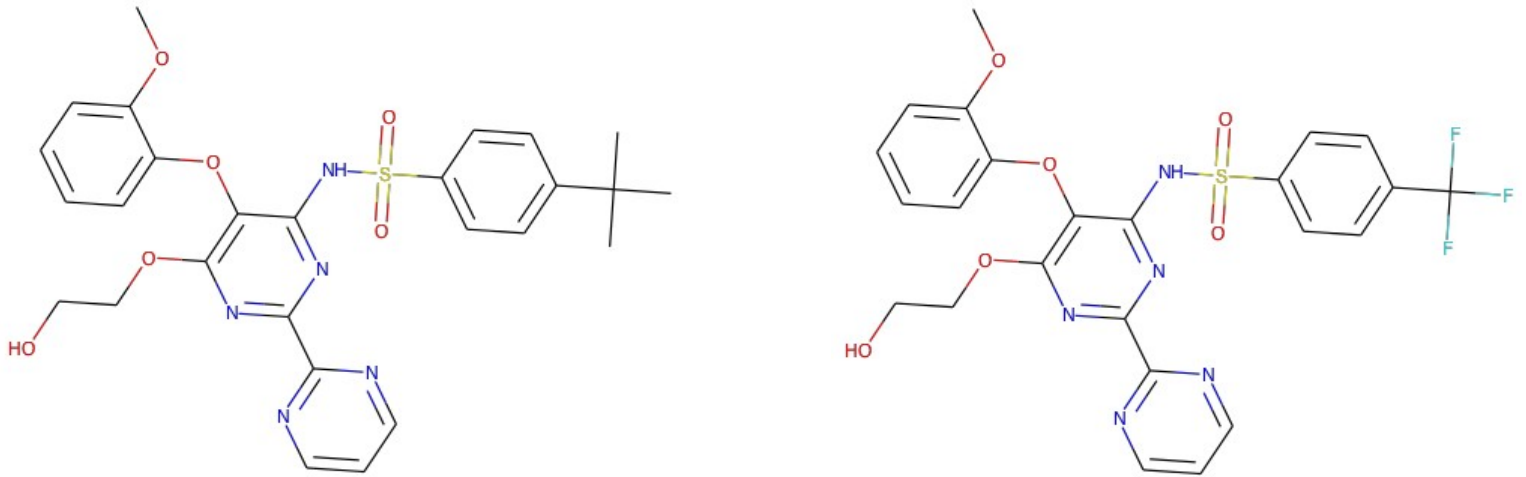
LUCID: supervised multidimensional optimization of drugs using Matched Molecular Pairs

Grégori Gerebtzoff, Ph.D.
Workshop on Chemical Information, September 12th 2014

Matched Molecular Pairs

A medicinal chemistry tool

"A Matched Molecular Pair is a pair of compounds that only differ by a single localized structural change" [1]



[1] Kenny, P. W. and Sadowski, J. (2005) Structure Modification in Chemical Databases, in Chemoinformatics in Drug Discovery (ed T. I. Oprea), Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG.

Computationally efficient algorithm to identify MMPs[1]

Systematic generation of all MMPs in chemical data sets

J. Chem. Inf. Model. **2010,** *50,* 339–348

Computationally Efficient Algorithm to Identify Matched Molecular Pairs (MMPs) in Large Data Sets

Jameed Hussain* and Ceara Rea

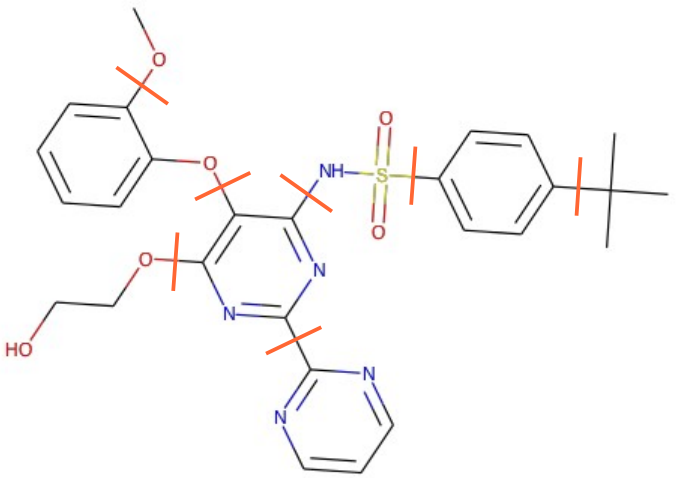
Computational & Structural Chemistry, GlaxoSmithKline, Medicines Research Centre, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2NY, U.K.

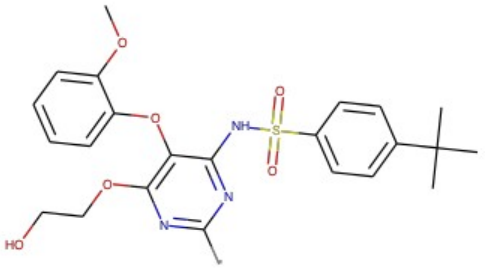
Received November 19, 2009

[1] J. Hussain & C. Rea, J. Chem. Inf. Model. 2010, 50 (3), 339

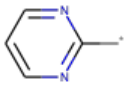
Computationally efficient algorithm to identify MMPs

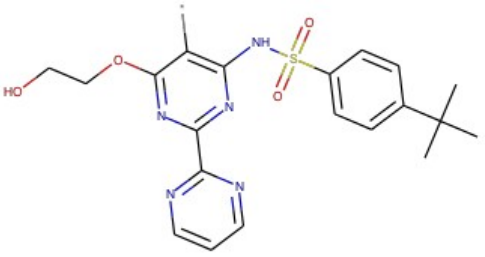
Systematic generation of all MMPs in chemical data sets



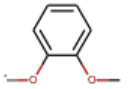


+



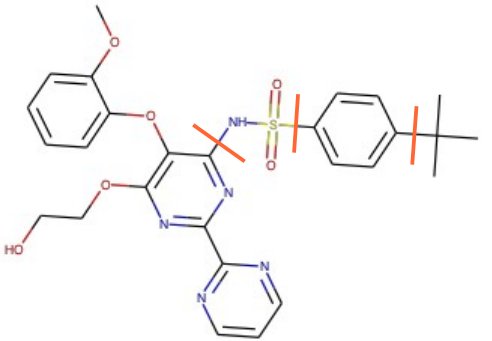


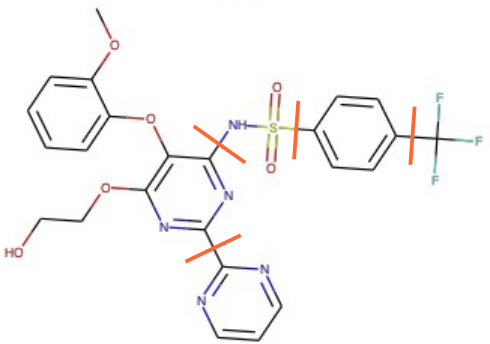
+

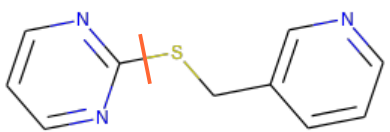


Computationally efficient algorithm to identify MMPs

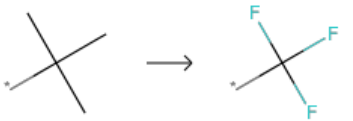
Systematic generation of all MMPs in chemical data sets

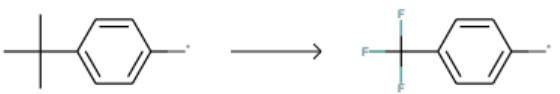


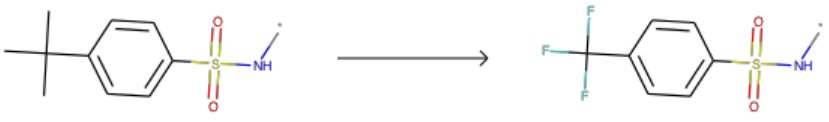


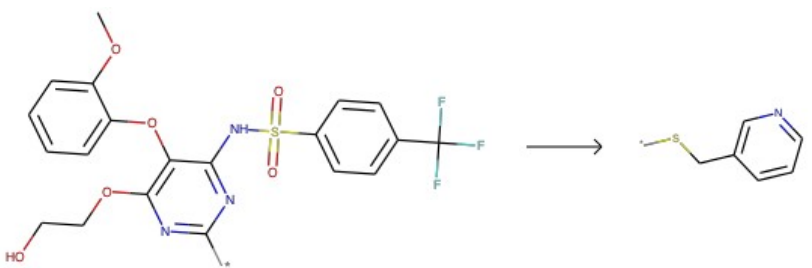


Tasuldine (expectorant)









Computationally efficient algorithm to identify MMPs

Usage and limitations of Hussain's method

Problems:

- The Hussain's algorithm identifies many MMPs and generates “too many” transformations
- In the original publication, 5.3 million MMPs identified (2.5 million unique transformations) for a dataset of 330'000 compounds generating 21.7 million fragmentations

Usage:

- Mainly identification and prediction of activity cliffs (SAR Maps)

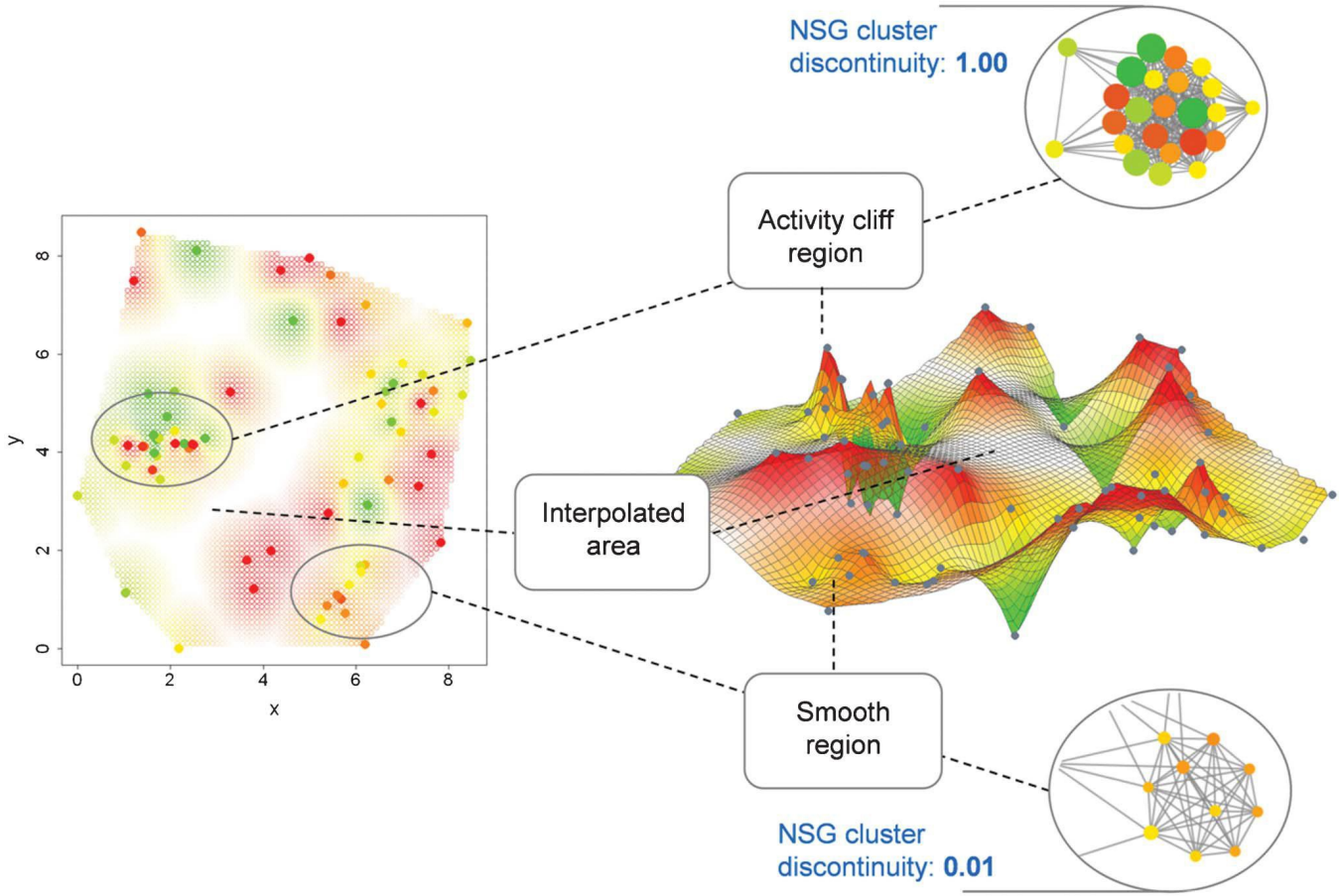


Illustration from: D. Stumpfe & J. Bajorath, RSC Adv. 2012, 2, 369

Can we use Hussain's method to help chemists designing better drugs?

Some thoughts on what I wanted to achieve

- The traditional MMP analysis approach allowed us to answer the question:
  “What is the effect of the replacement of side chain X to side chain Y on various properties (lipophilicity, membrane permeability, …)”
- What scientists expected to get:
  Easy guidelines = a catalog of transformations so solve MedChem issues
- New paradigm: LUCID
  Instead of searching for the effect of specific replacements, LUCID helps the chemists optimizing various properties by suggesting potentially beneficial replacements

LUCID is based on the Hussain's method, but suggests a lower number of relevant MMPs

Using Hussain's method to extend traditional MMP analysis

Supervised multidimensional optimization of drugs

- Define the size of the replacement
- Define the local similarity of the core
- Select required features in the replacement (aromatic ring, …)
- List of properties (physchem, safety, …) to optimize
- Define the site(s) of transformation

LUCID
\'lü-səd\

“Very clear and easy to understand”
“Able to think clearly”

Acknowledgments

Roche

Stefanie Bendels (original ComPair-pC implementation)

Manfred Kansy

Wolfgang Guba

Olivier Roche

...and all the LUCID users

Others

Greg Landrum (RDKit)

Jameed Hussain

LUCID web interface

Interactive visualization of Chemical Space and MMPs

Dataset

Current set (updated):

Sutherland_DHFR

The dataset contains 756 molecules

Work with one SRN / smiles / structure:

44-10 ...

Search & display chemical space

Search & display MMPs

Show only compact modules

Use only pairs of the current dataset

Submit compound No pictures

Chemical Space

Sort data by:

Don't sort

P carinii pIC50

Rat Liver pIC50

T gondii pIC50

Matched Molecular Pairs

Settings:

Number of atoms replaced: **-4 ≤ Δn ≤ 4**

Consider local environment: **off (all pairs)**

Replacement contains:

any ring, aromatic ring, anything

Sort replacements by:

number of results per MMP, similarity

Desired effect:

P carinii pIC50

8.3 ≤ ΔP carinii pIC50 ≤ 10.4

Rat Liver pIC50

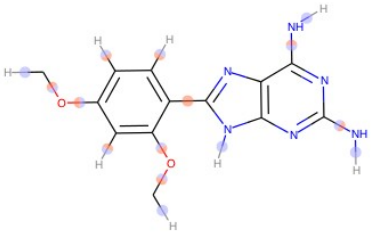
8.7 ≤ ΔRat Liver pIC50 ≤ 10.9

T gondii pIC50

8.6 ≤ ΔT gondii pIC50 ≤ 10.9

Help (download the PowerPoint file)

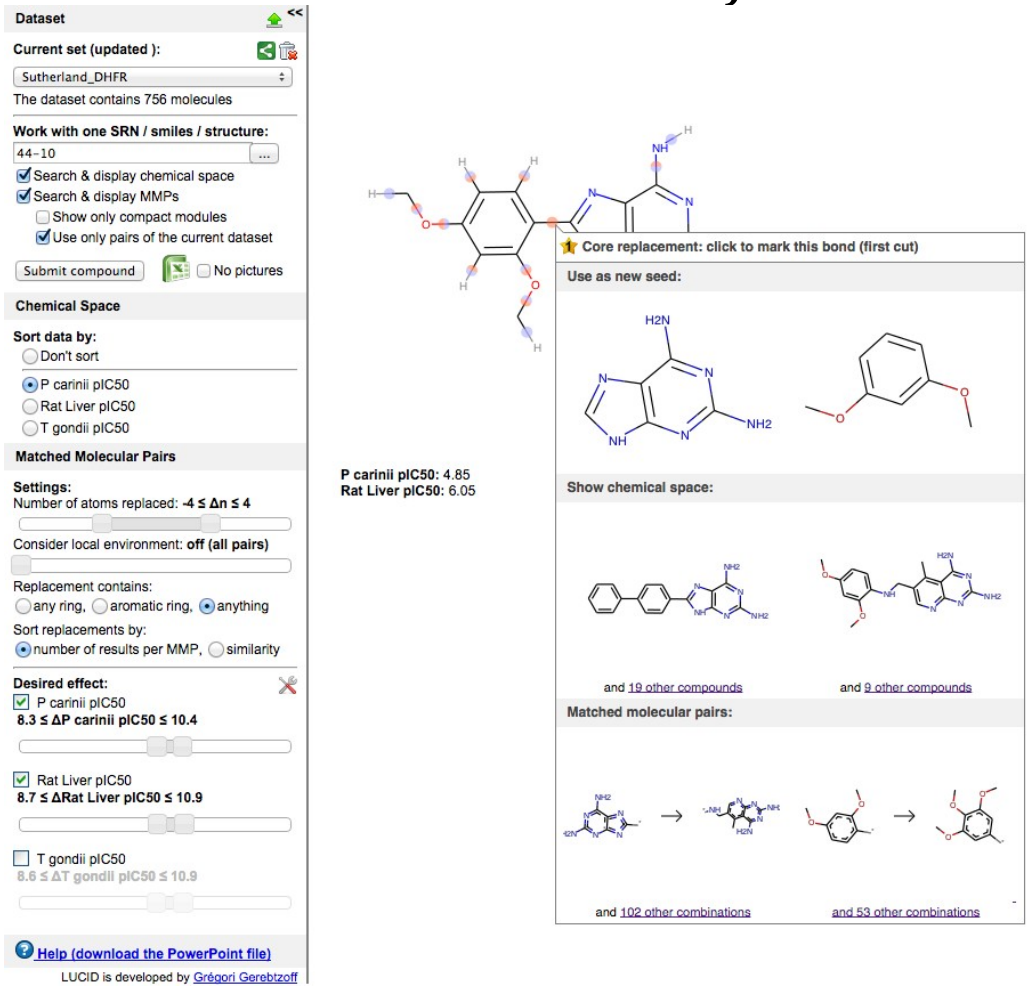
LUCID is developed by Grégori Gerebtzoff



P carinii pIC50: 4.85
Rat Liver pIC50: 6.05

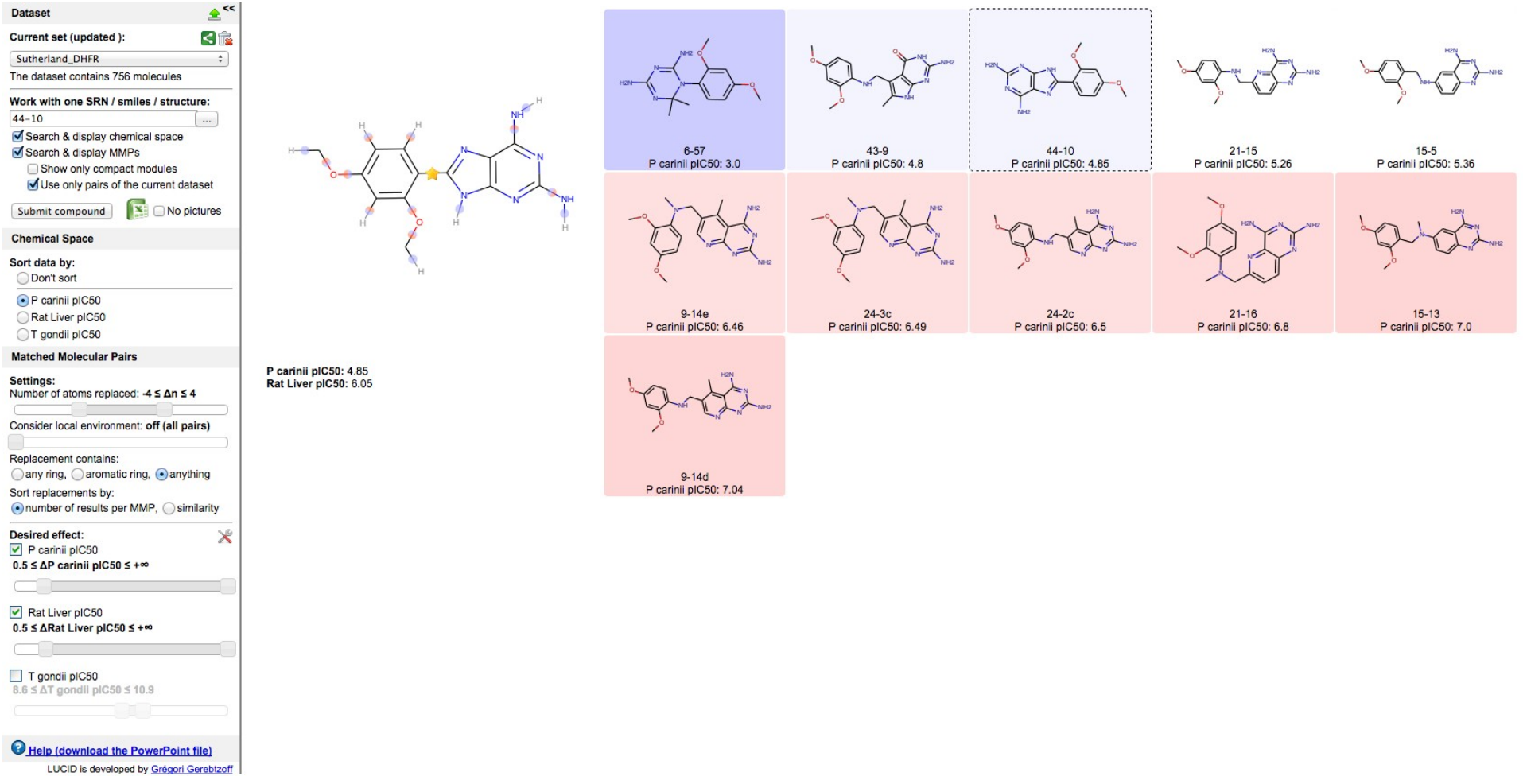
LUCID web interface

Interactive selection of the site of transformation



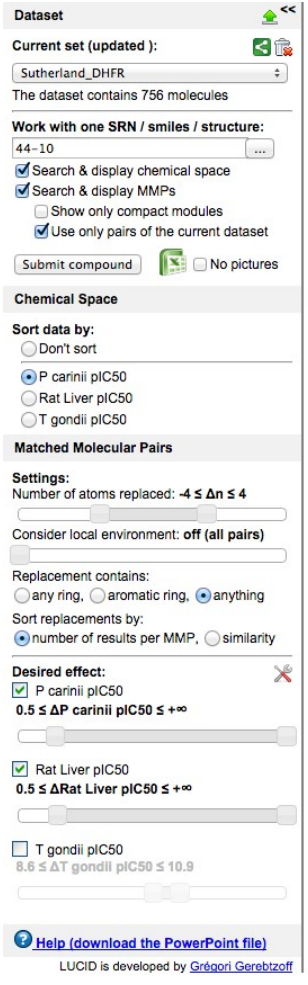
LUCID web interface

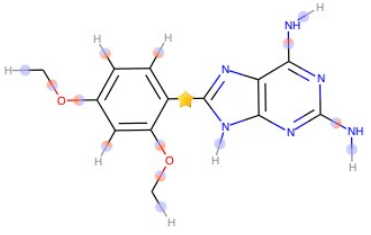
Chemical space around one constant fragment, colored by the



LUCID web interface

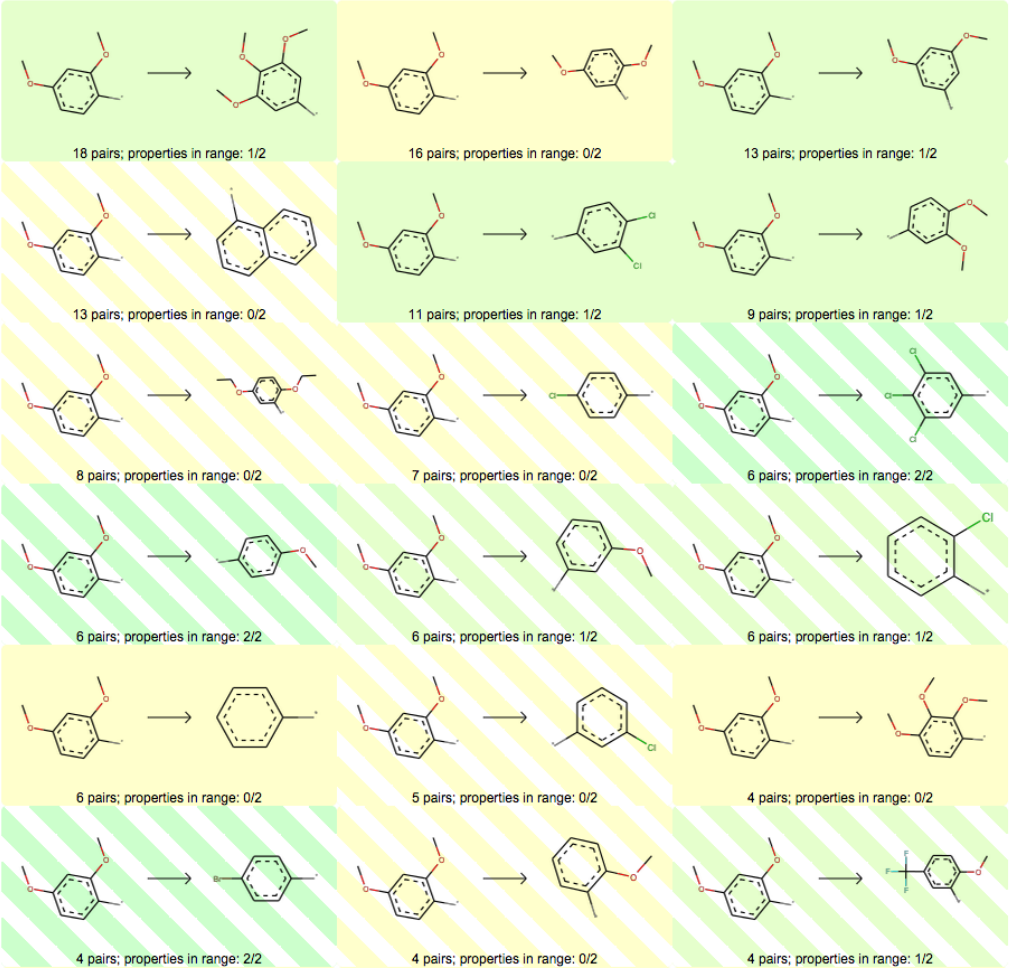
Transformation at the selected site, colored by the number of "good"





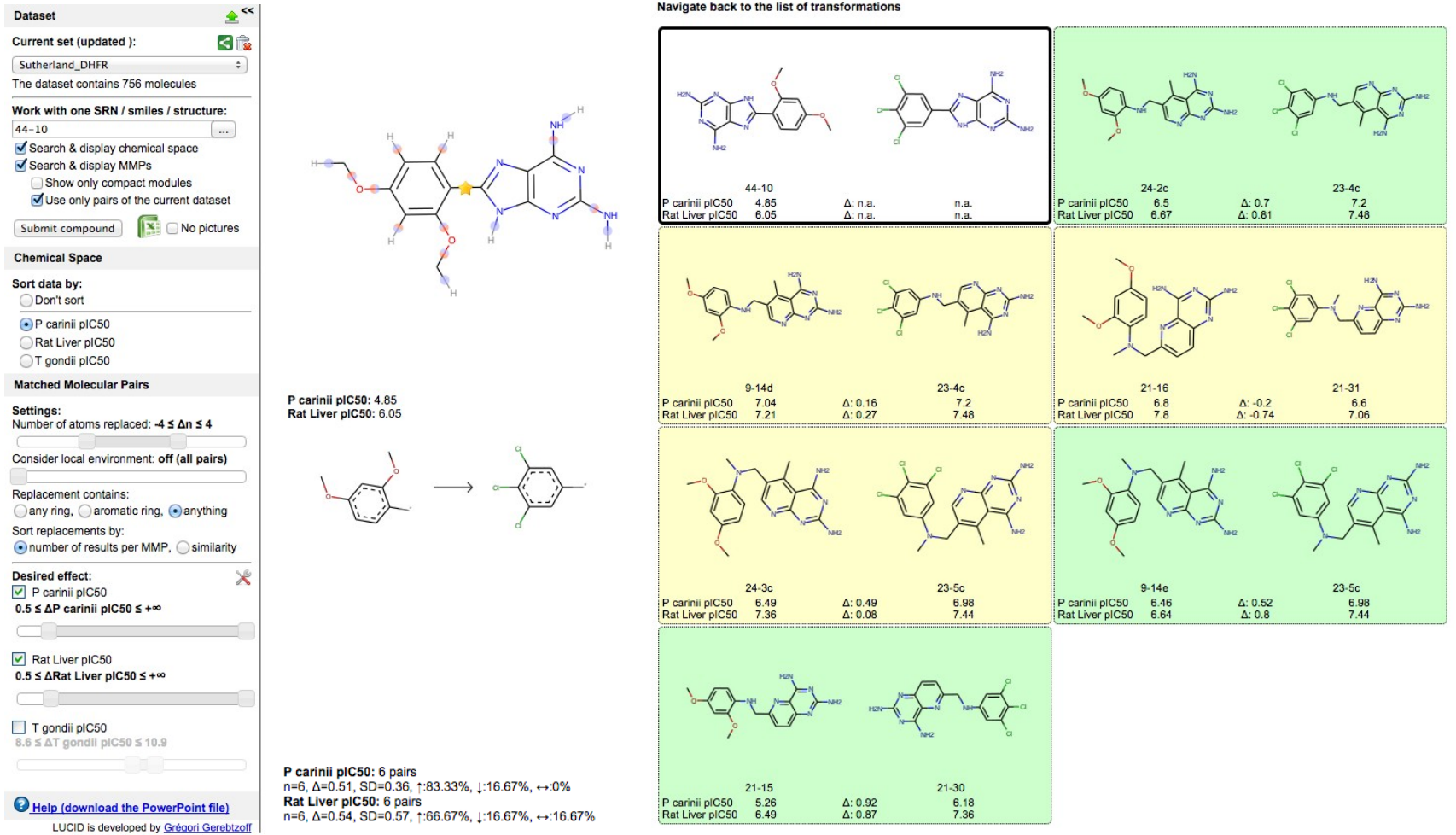
P carinii pIC50: 4.85
Rat Liver pIC50: 6.05

P carinii pIC50: 6 pairs
n=6, Δ=0.51, SD=0.36, ↑:83.33%, ↓:16.67%, ↔:0%
Rat Liver pIC50: 6 pairs
n=6, Δ=0.54, SD=0.57, ↑:66.67%, ↓:16.67%, ↔:16.67%



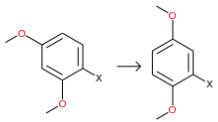
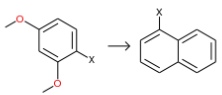
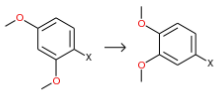
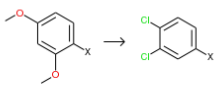
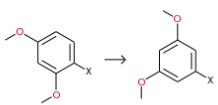
LUCID web interface

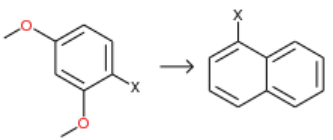
Pairs of compounds (MMPs) corresponding to the selected



Excel export

Transformations (left) and MMPs of one particular transformation

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1			Local Similarity		user_Rat_Liver_pIC50							user_P_carinii_pIC50						
2	Transformation	Transform	Size	Strength	N	Median	StdErr	StdDev	Increase	Decrease	Neutral	N	Median	StdErr	StdDev	Increase	Decrease	Neutral
3		44-7	5	0	8	-0.3	0.14	0.4	2	5	1	8	0.021	0.15	0.43	3	3	2
4		Nc1nc2[nH	5	0	6	0.6	0.2	0.49	5	1	0	6	0.69	0.2	0.5	5	1	0
5		44-9	5	0	5	0.77	0.51	1.1	4	1	0							
6		44-11	5	0	5	0.02	0.52	1.2	2	2	1							
7		44-8	5	0	5	0.15	0.37	0.83	3	1	1							

	A	B	C	D	E	F	G	H	I
1				user_Rat_Liver_pIC50			user_P_carinii_pIC50		
2	Compound1	Compound2	Local Simi	Value1	Value2	Delta	Value1	Value2	Delta
3	44-10	Nc1nc2[nH]c(-c3cccc	1	6.04576	n.a.	n.a.	4.85387	n.a.	n.a.
4	24-3C	9-14M	0.17	7.35655	6.85387	-0.50268	6.49485	6.82391	0.32906
5	24-2C	24-5A	0.17	6.66959	7.52871	0.85912	6.50031	6.24185	-0.25846
6	21-16	21-38	0.17	7.79588	8.13668	0.3408	6.79588	7.39794	0.60206
7	15-13	15-15	0.17	7.36653	7.76955	0.40302	7	7.76955	0.76955
8	21-15	21-36	0.17	6.49485	7.39794	0.90309	5.25964	6.63827	1.37863
9	15-5	15-10	0.17	5.92082	6.72125	0.80043	5.35655	6.14267	0.78612