Chemonaut

on-line compound sourcing platform

Anna Rzepiela

Date 20 May 2010
ChemAxon UGM

PYXIS discovery

Pyxis Discovery

- **Chemistry based lead discovery services**
- **Effective application of computational chemistry tools for faster lead identification using fewer compounds**
- **3 focus areas**
  - Chemistry-based lead discovery services
  - Smart Libraries
  - Chemonaut

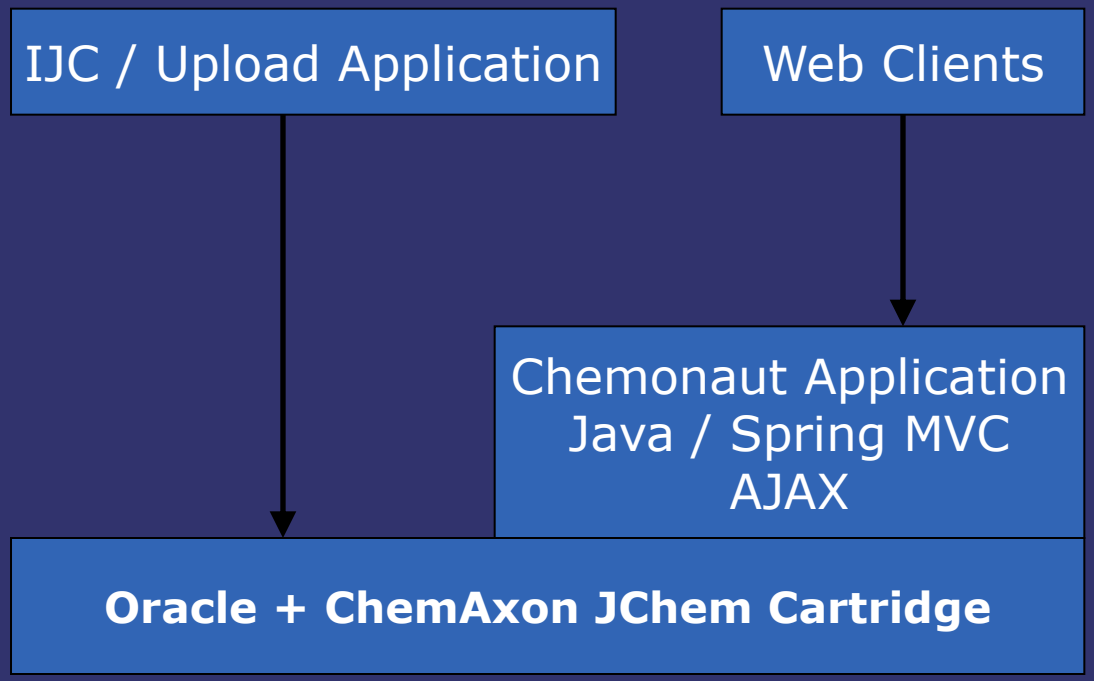
PYXIS discovery

Chemonaut

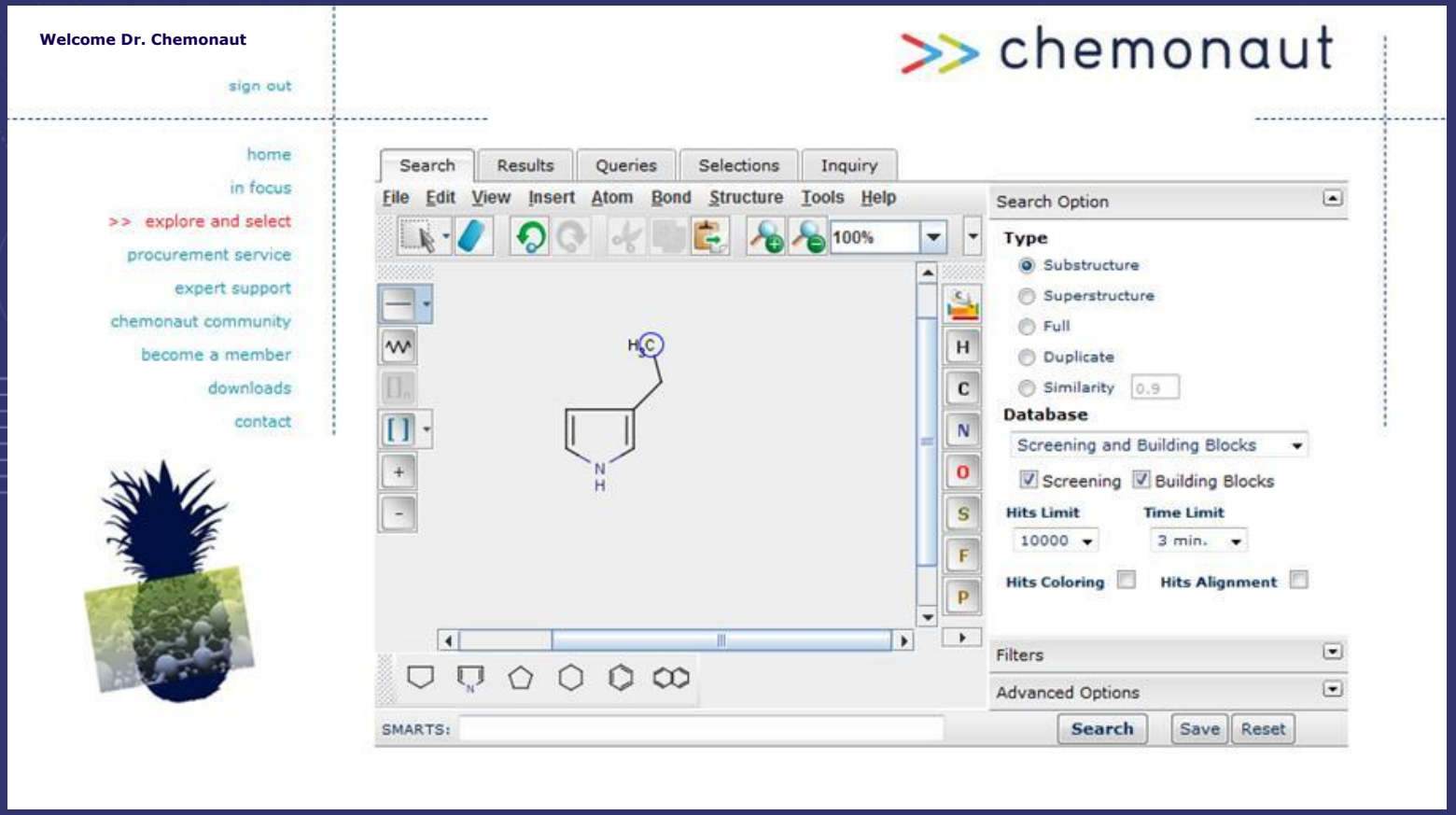
- Database of screening compounds and building blocks from reliable suppliers
- Standardized structures and data
- Leadlike filtering
- Physico-chemical properties calculated

PYXIS discovery

Chemonaut: application

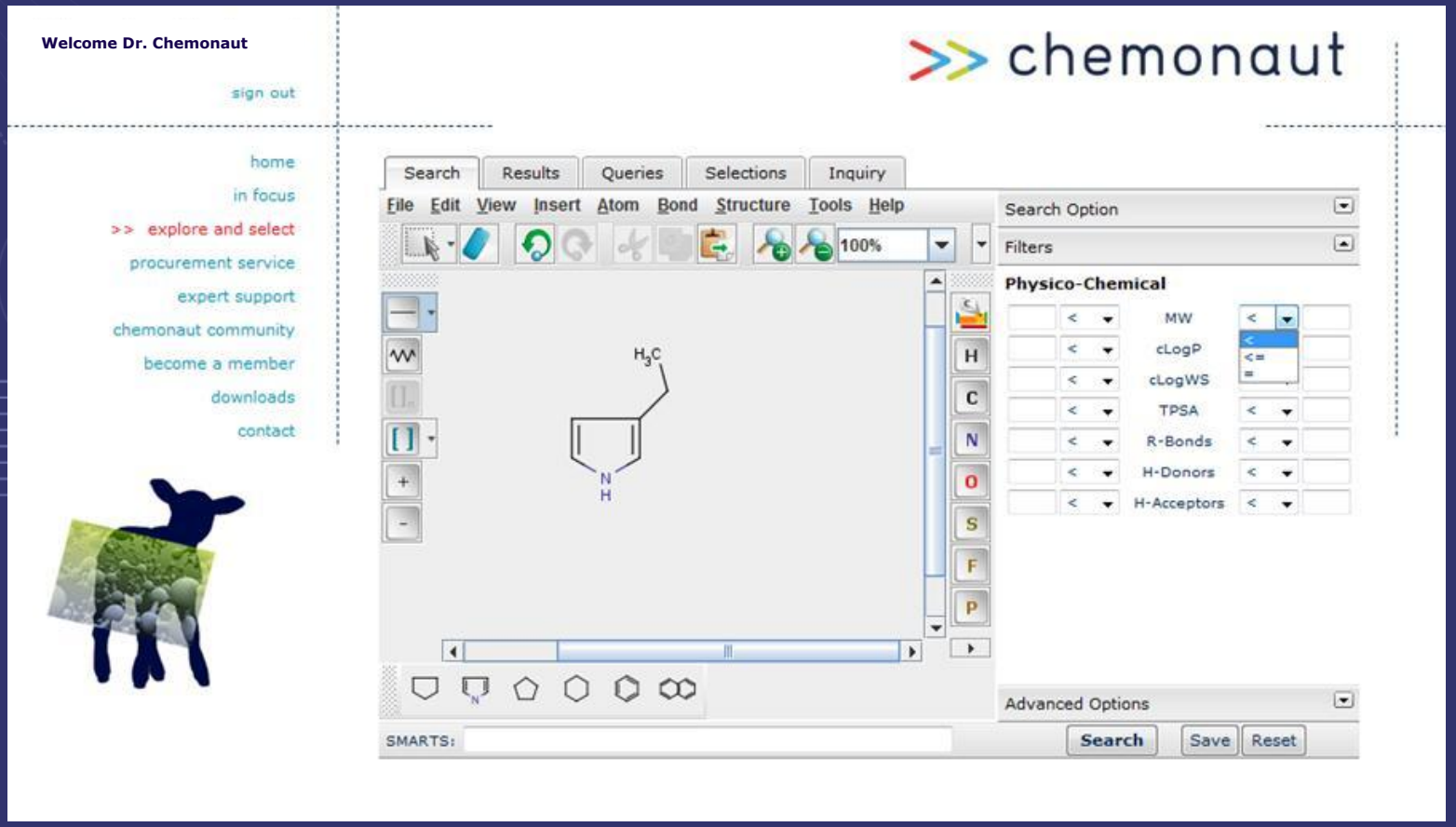


Chemonaut: web view



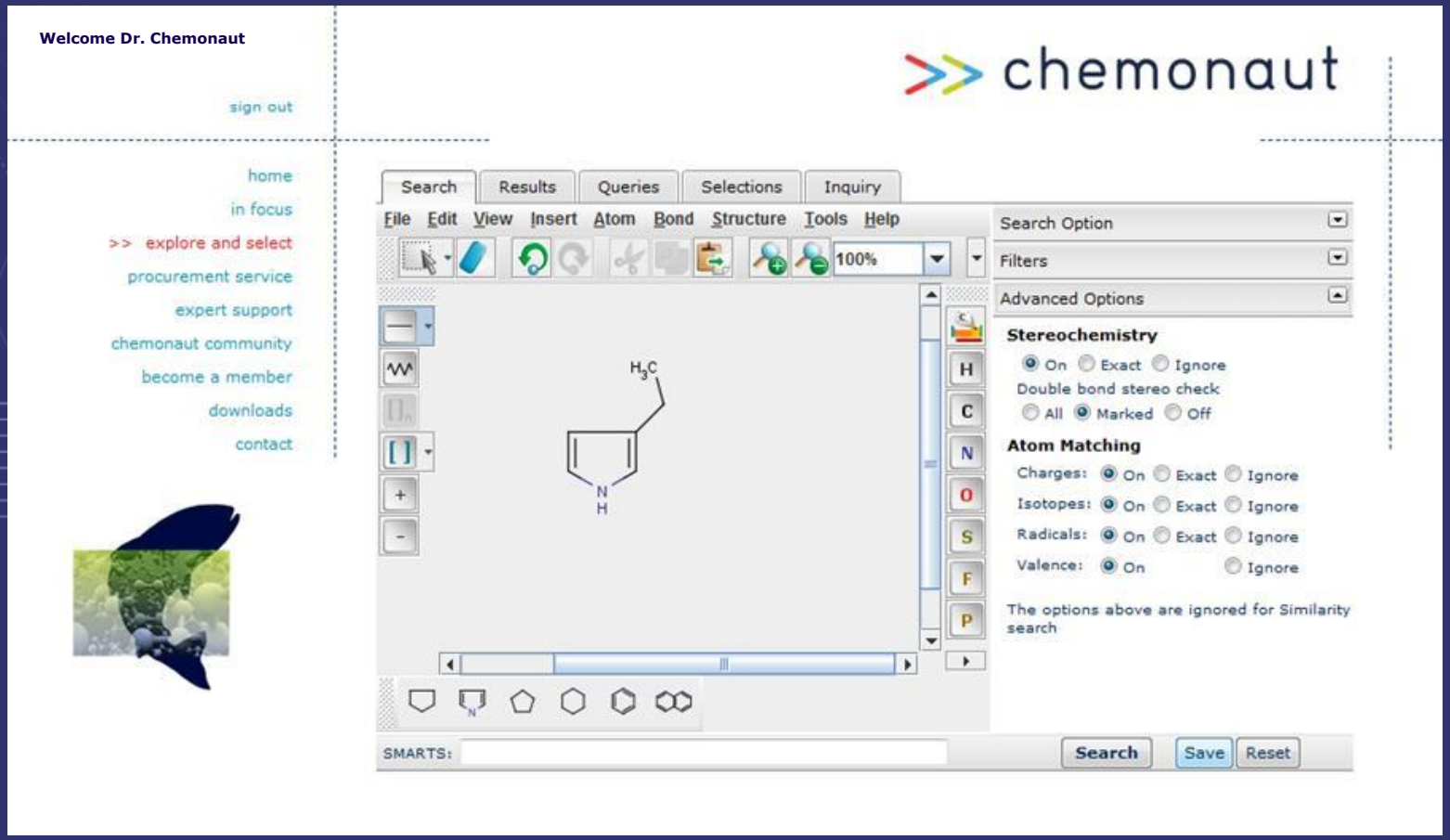
PYXIS discovery

Chemonaut: web view



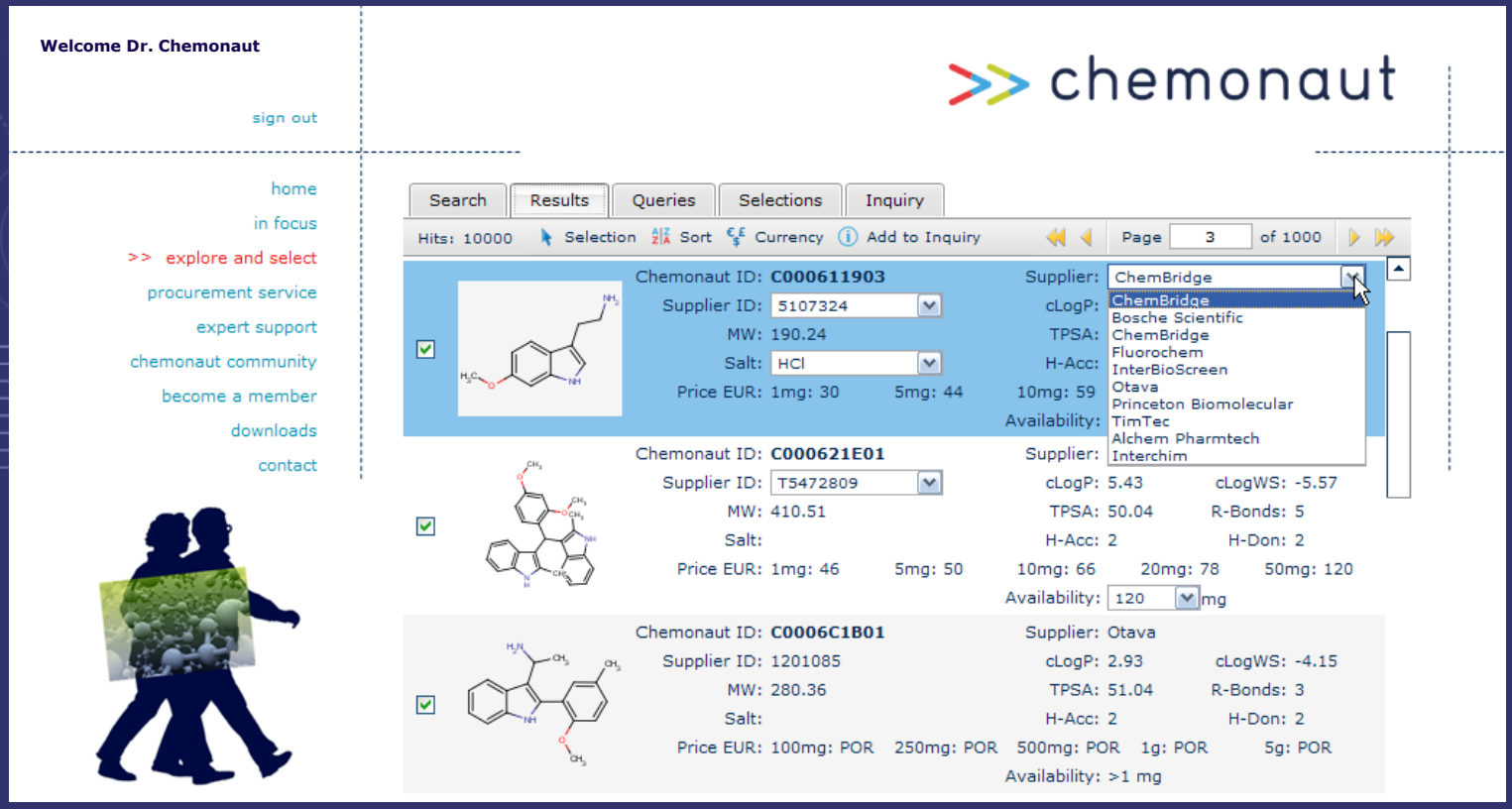
PYXIS discovery

Chemonaut: web view

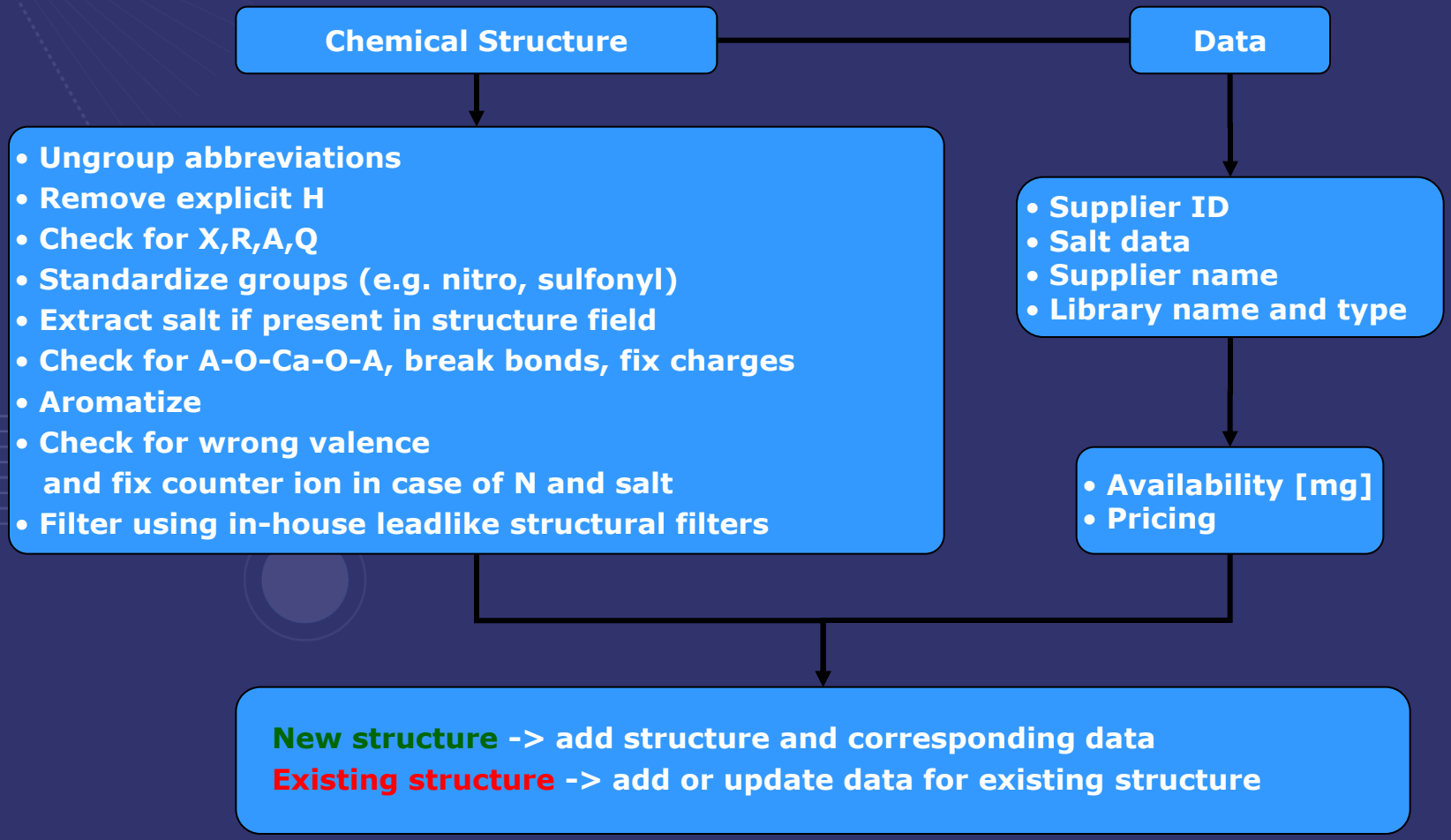


PYXIS discovery

Chemonaut: web view



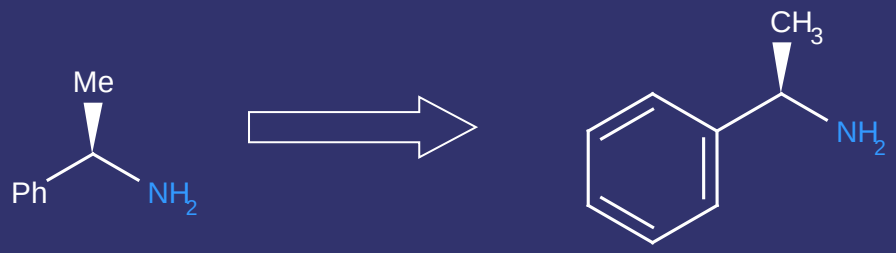
Chemonaut: procedure



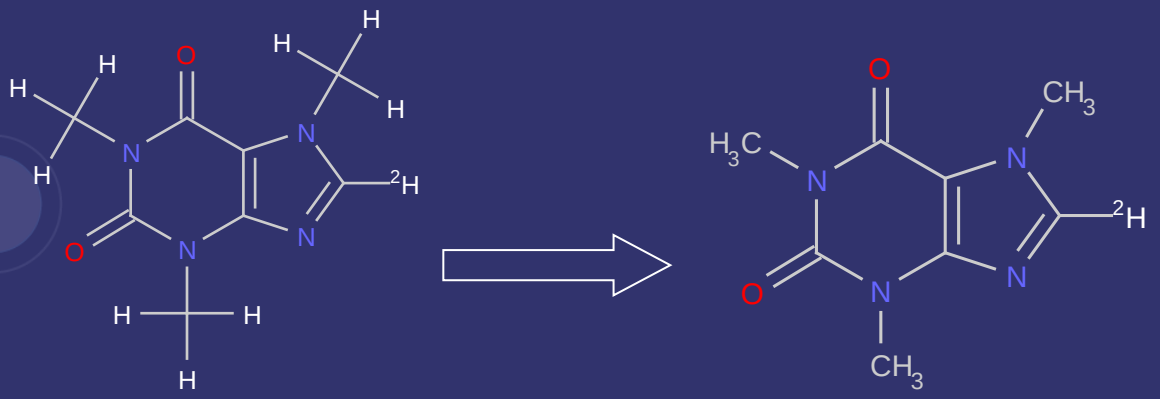
PYXIS discovery

Structure import: standardize

- **Ungroup abbreviations**



- **Remove explicit Hydrogen atoms**



Converts explicit hydrogens to implicit (removes hydrogen atoms from the molecule graph). Charged, radical, mapped, isotope or wedged hydrogens are not removed.

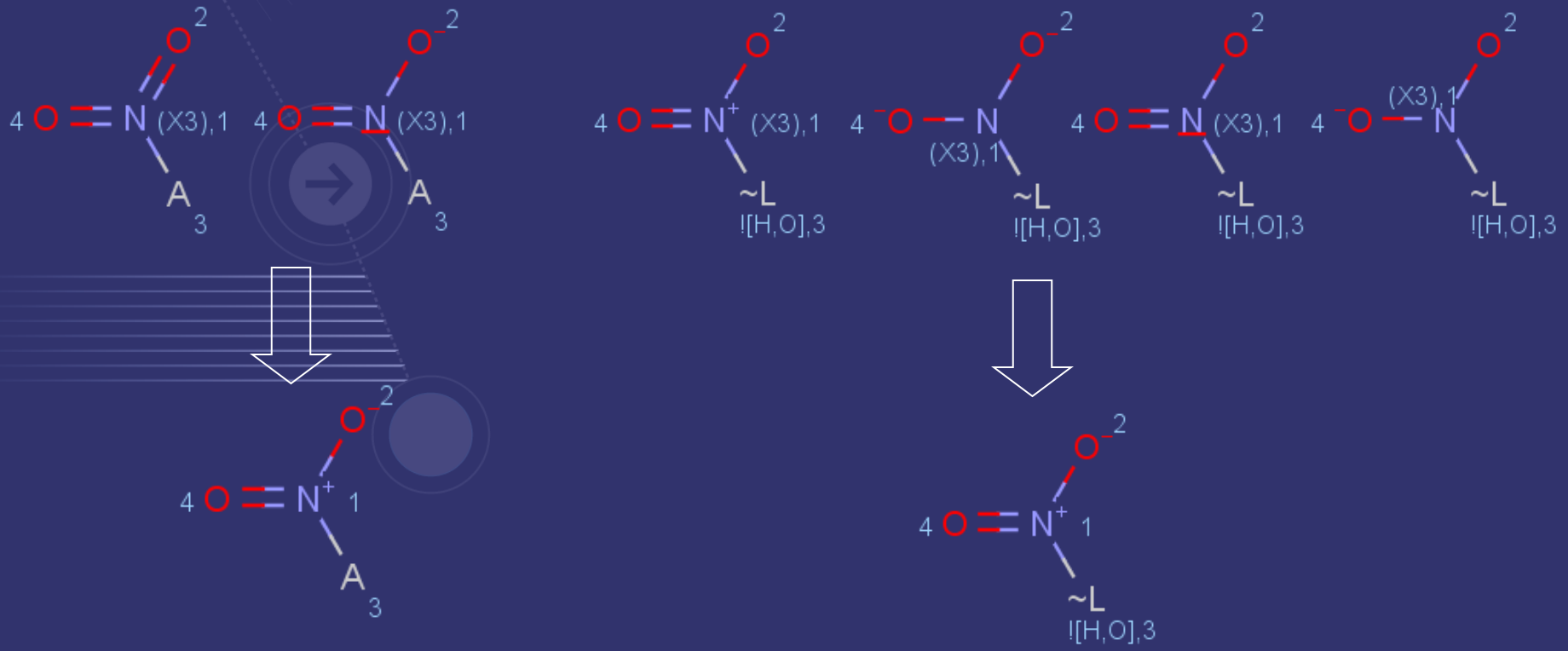
Structure import: wash

- Check for X,R,A,Q and discard if found

PYXIS discovery

Structure import: standardize

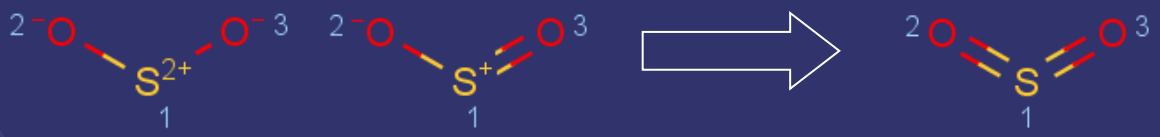
- **Standardize groups, e.g. nitro group**



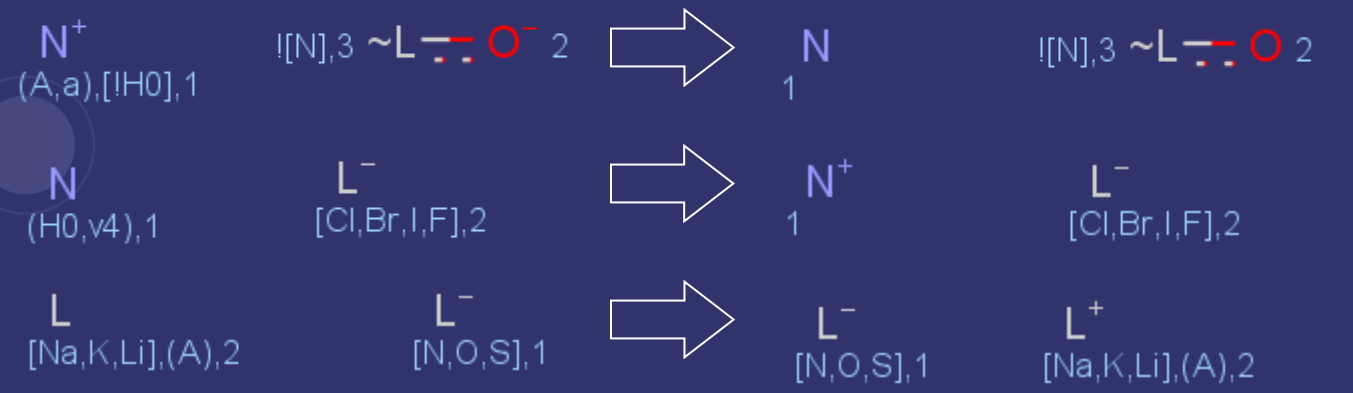
PYXIS discovery

Structure import: standardize

- **Standardize groups, e.g. sulfonyl**



- **Fix charges, e.g.: match charge on structure and salt, protonate acids, deprotonate bases:**

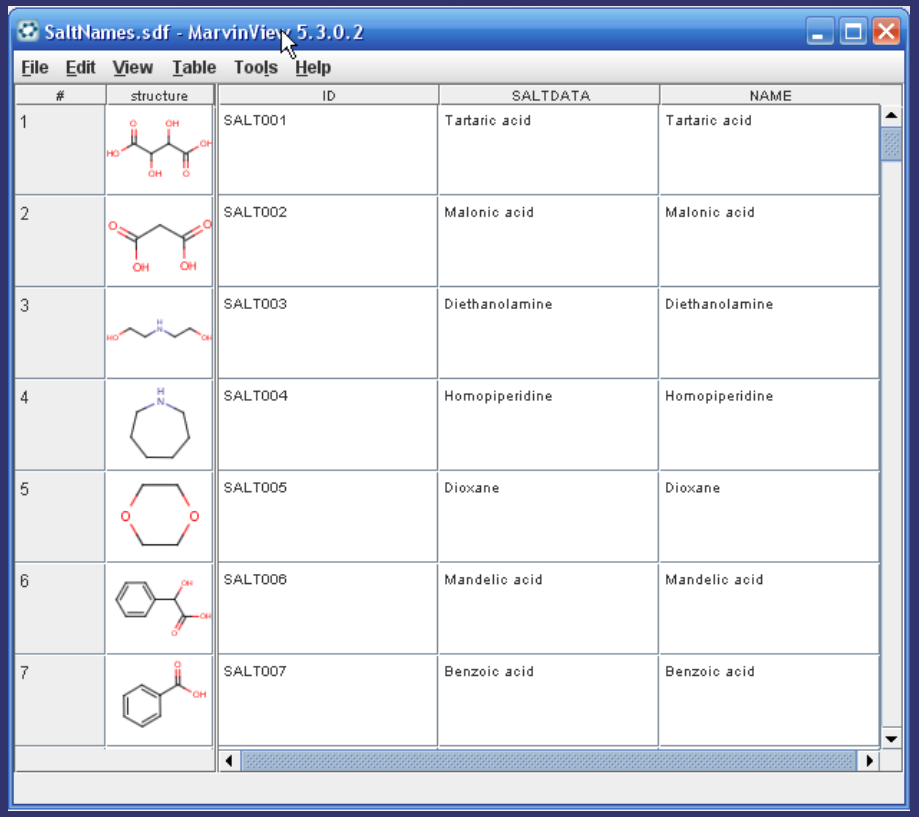


PYXIS discovery

Structure import: wash

- **Move salt data from structure field to Salt data field**

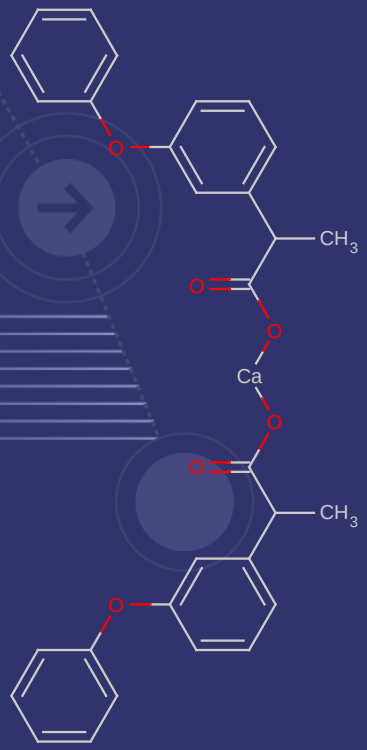
for records with more than one structure in the structure field check if exists in the salt table store salt and ratio in salt field

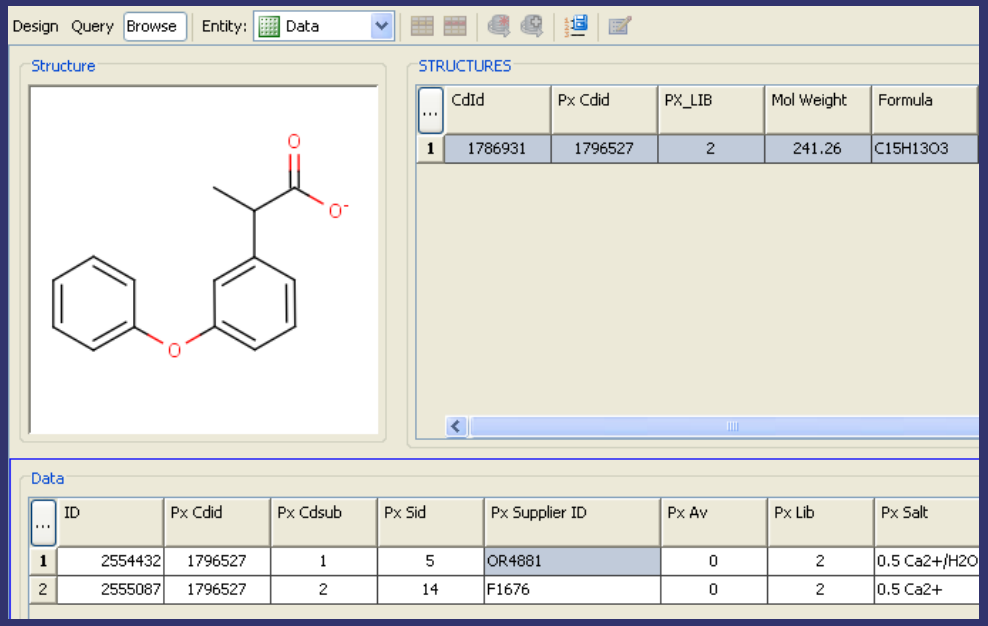


PYXIS discovery

Structure import: standardize

- Check for A-O-Ca-O-A, break bonds, fix charges

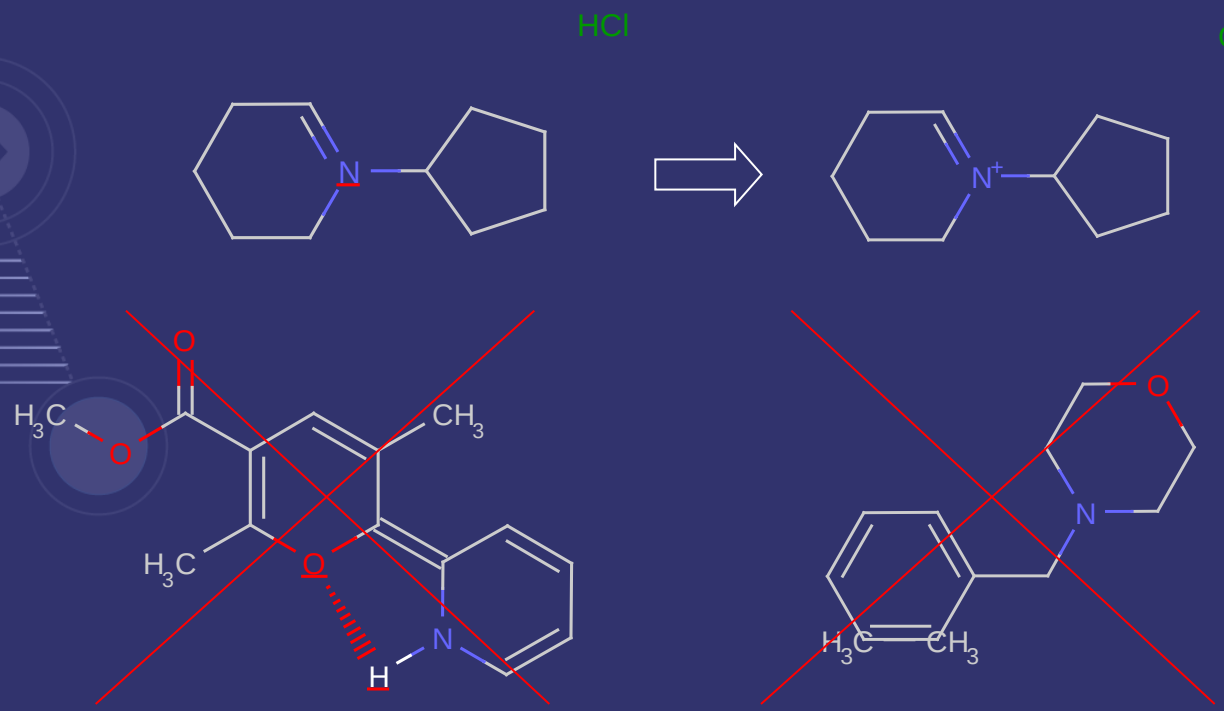




PYXIS discovery

Structure import: standardize

- **Check for wrong valence and fix counter ion in case of N and salt, discard the rest**



PYXIS discovery

Chemonaut: leadlike filtering

remove the false positives, the covalent-acting electrophiles, the 'suicide inhibitors', the 'magic bullets', the 'warheads', etc., as well as the dyes, fertilizers, photographic chemicals, etc.

In total, a set of over 80 structural filters

1. Rishton, G.M., Reactive compounds and in vitro false positives in HTS, Drug Discovery Today, 2 (1997) 382–384.
2. Rishton, G.M., Nonleadlikeness and leadlikeness in biochemical screening, Drug Discovery Today, 8 (2003) 86–96.
3. Rishton, G.M., Failure and success in modern drug discovery: guiding principles in the establishment of high probability of success drug discovery organizations, Medicinal Chemistry, 1 (2005) 519–527.
4. Olah, M.M. *et al.*, Strategies for Compound Selection, Current Drug Discovery Technologies, 1 (2004), 211-220.
5. Huth, J.R. *et al.*, ALARM NMR: A rapid and robust experimental method to detect reactive false positives in biochemical screens, J. Am. Chem. Soc., 127 (2005) 217–224.
6. McGovern, S.L. *et al.*, A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening, J.Med. Chem., 45 (2002) 1712–1722.
7. Seidler, J. *et al.*, Identification and prediction of promiscuous aggregating inhibitors among known drugs, J. Med. Chem., 46 (2003) 4477–4486.
8. Roche, O. *et al.*, Development of a virtual screening method for identification of 'frequent hitters' in compound libraries, J. Med. Chem., 45 (2002) 137–142.

PYXIS discovery

Chemonaut: leadlike filtering

1. Chemically reactive (sub)structures

- …aldehydes, isocyanates, hydrazones and hydrazides, maleimides…
  …[C;H1](*)=O, N=[C,S]=[N,O,S]…

2. Biologically unstable, cytotoxic (sub)structures and frequent hitters

- …thiols, quinones, thioketones, aminothiophenes, diphenyl diazenes…
  …[S;H1]*, O=C1C=CC=CC1=O…

3. Chemical structures difficult in follow up

- …structures containing: x CF2 groups, nitrate and sulfate groups, >2 nitro groups, no hetero-atoms…
  …FC(F)(*)C(F)(F)C(F)(F)*, [#8-][N+](=O)O!@-*…

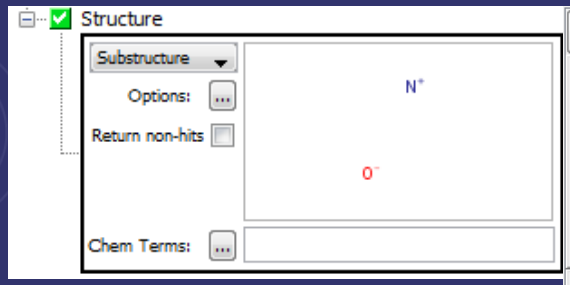
Chemonaut: future developments

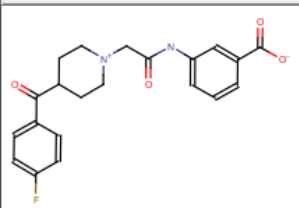
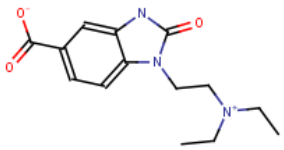
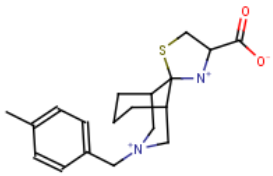
- Tautomer enumeration
- pKa prediction and enumeration of ionized macrospecies
- Conformer enumeration and 3D substructure search

PYXIS discovery

pKa prediction and enumeration of ionized macrospecies

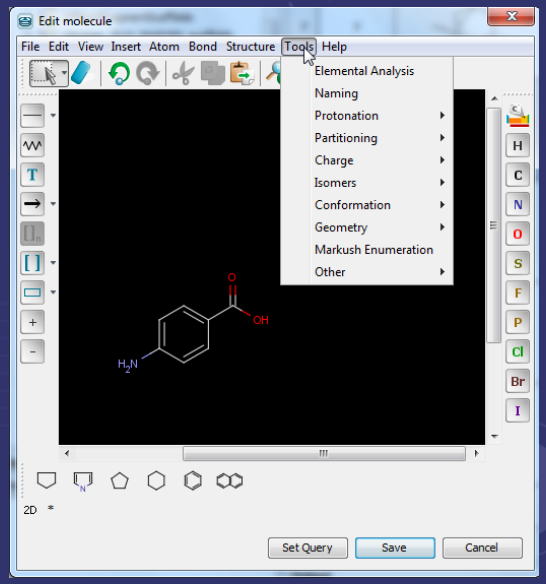
- advanced search options
- structure preparation for virtual screening

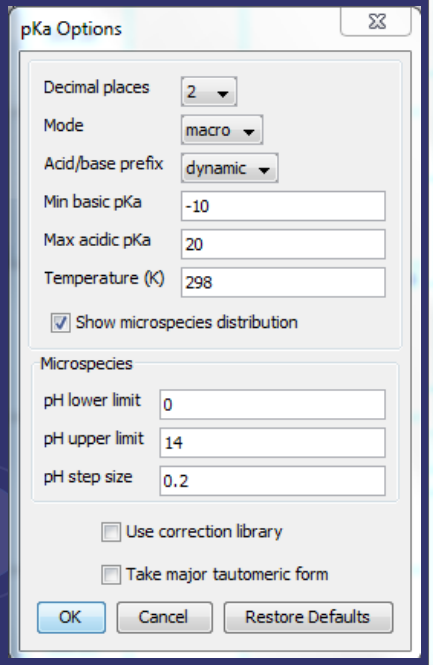


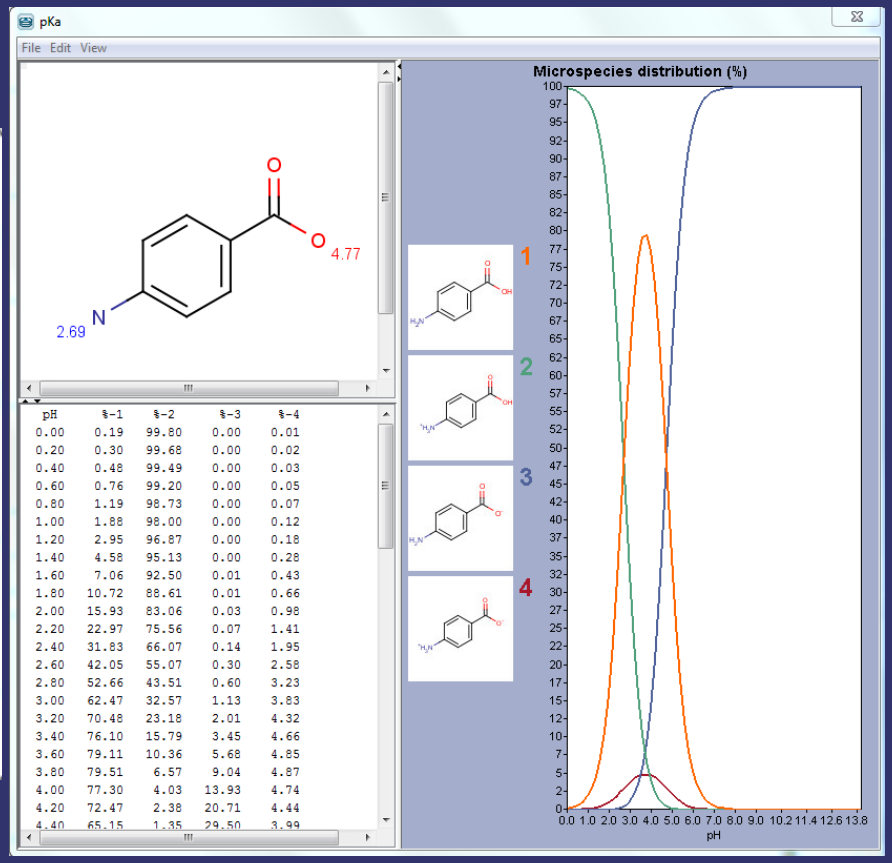
	CdId	Structure	Mol Weight	Formula	IDNUMBER
1	3,512		384.40	C21H21FN2O4	PYX 00004098
2	4,265		277.32	C14H19N3O3	PYX 00005416
3	8,841		347.49	C19H27N2O2S	PYX 00011410

PYXIS discovery

Protomers: ChemAxon pKa predictor







cxcalc pka -i -10 –x 20 –a 3 –b 3 Input.sdf >Output.tbl

PYXIS discovery

Protomers: ChemAxon pKa predictor

- Good agreement with experimental values in terms of RMSE
- Good detection success rate
- Predicts many protomers

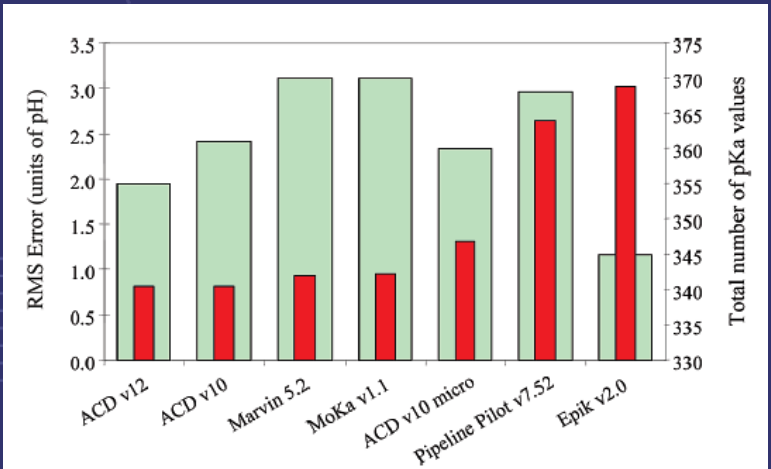


Figure 2. RMSE for the 7 different approaches examined (red/dark bars), superimposed on the total number of pK_a values predicted by each approach (green/light bars).

	measured	batch pKa (standard settings)	batch pKa (-i -2 -x 12)
basic groups	135	249	230
acidic groups	7	43	47

Manchester *et al.* J. Chem. Inf. Model. 2010 xxx xxx

Lee *et al.* J. Chem. Inf. Model. 2009 49 2013

Liao *et al.* J. Chem. Inf. Model. 2009 49 2801

Protomers: ionized forms enumeration

- "cxcalc msdistr -H 7.4 Input.sdf > Output.sdf"
- command line to write the microspecies distribution of an "Input.sdf" at given pH into an "Output.sdf"
- "Output.sdf" contains protonated microspecies at pH 7.4 and [%] distribution value of microspecies at this pH
- no –i, -x, -a, -b nor pH range or [%] distribution min
- ~30 protomers per molecule (leadlike molecules)

PYXIS discovery

Chemonaut

Screening - 3,916,796 structures from which 2,539,616 are UNIQUE
19 suppliers

Building Blocks - 622,279 structures from which 452,770 are UNIQUE
46 suppliers

Special - 887,156 structures from which 591,336 are UNIQUE
20 suppliers

Overall – 5,426,231 structures from which 3,151,539 are UNIQUE
51 suppliers

Virtual - 11 million structures
9 suppliers

PYXIS discovery

Thank you

PYXIS discovery