

Naming tools

Daniel Bonniot & Alex Allardyce



ChemAxon

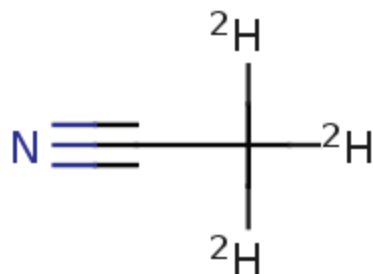
Products

- Structure to Name
- Name to Structure
- Document extractor

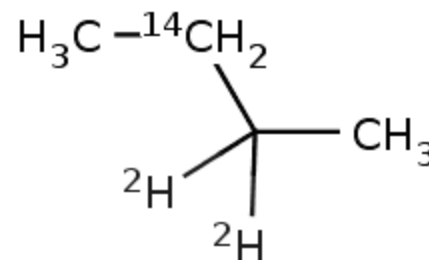


Structure to Name

- Support for Isotopes:



(²H₃)acetonitrile



(3-¹⁴C,2,2-²H₂)butane

- Generates more traditional names (optional)
- Various fixes and smaller improvements



Name to Structure

Continuous improvement in 5.3.x of:

- Coverage (how many names we understand correctly)
- Correction (few errors)

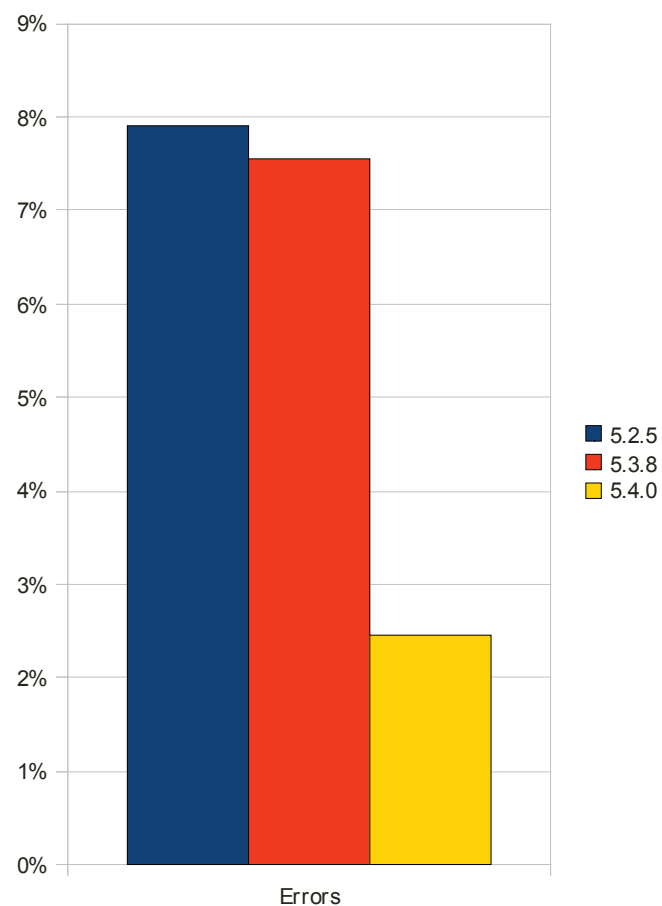
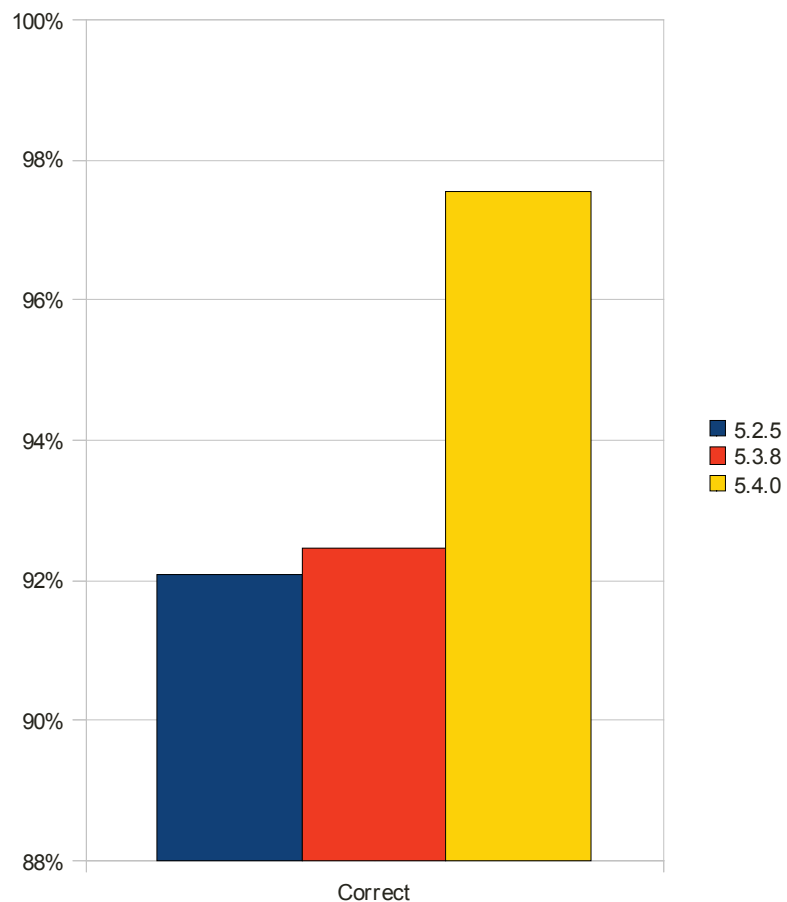
5.4.0 will include a major upgrade:

- Correction++
- Coverage+



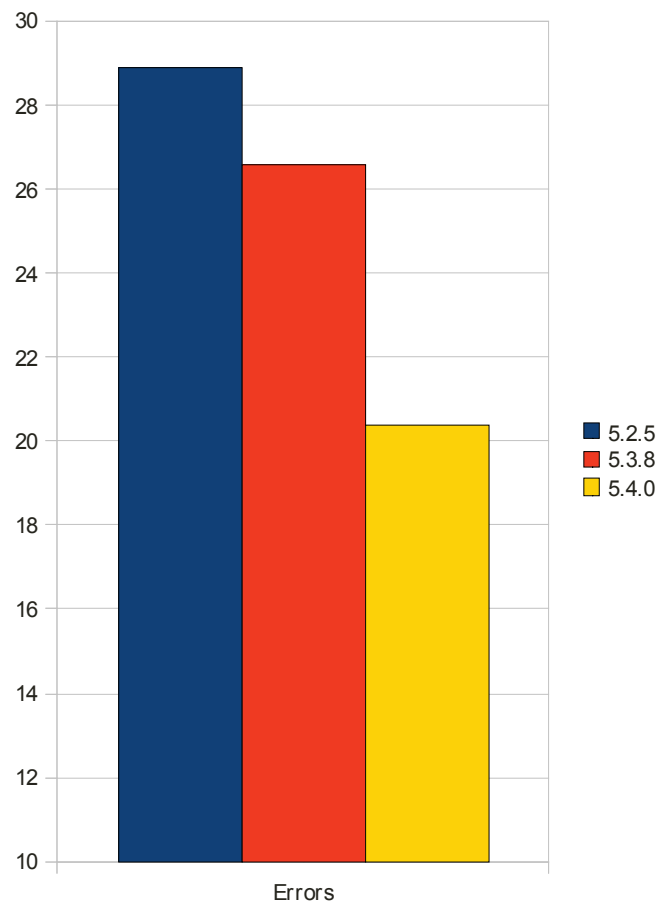
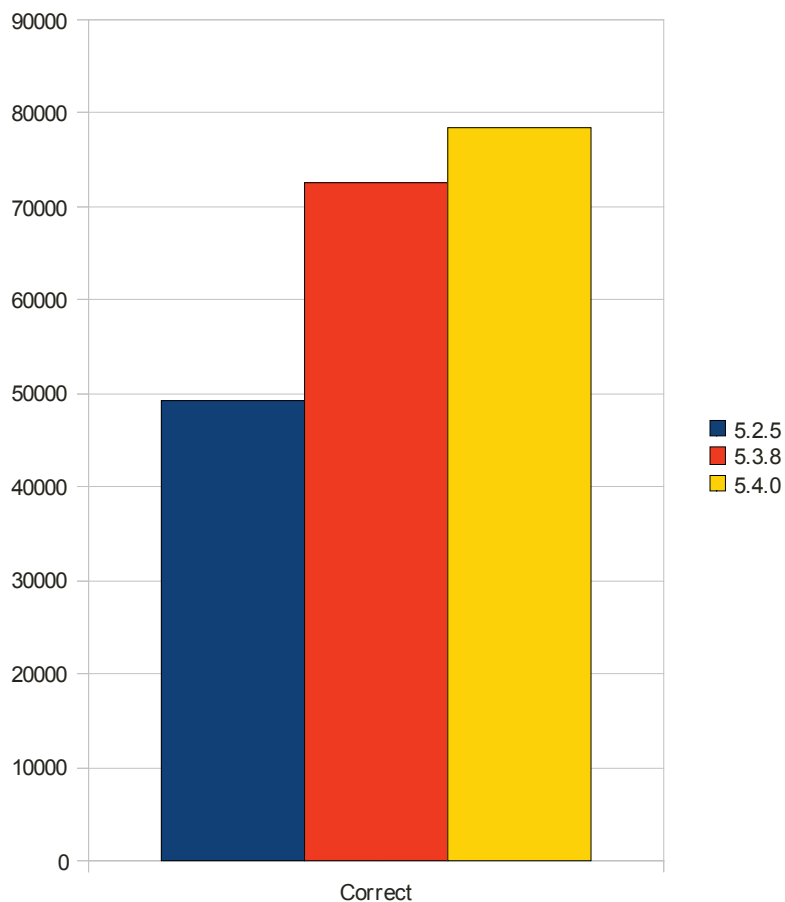
Name to Structure improvements

Systematic names (easier)



Name to Structure improvements

Various names, mixed nomenclatures (hard)



Document to Structure

- Extracts names from text documents
- Returns the discovered structures and their location
- New in 5.4: extraction from **PDF** documents



Structure to Name: incremental improvements, mature

Name to Structure: improving steadily, major upgrade in 5.4

Document to Structure: improves with n2s, PDF support

Try them, get support, ask questions!

All very nice but how can I use it

chemicalize >



chemicalize.org : What's it for?

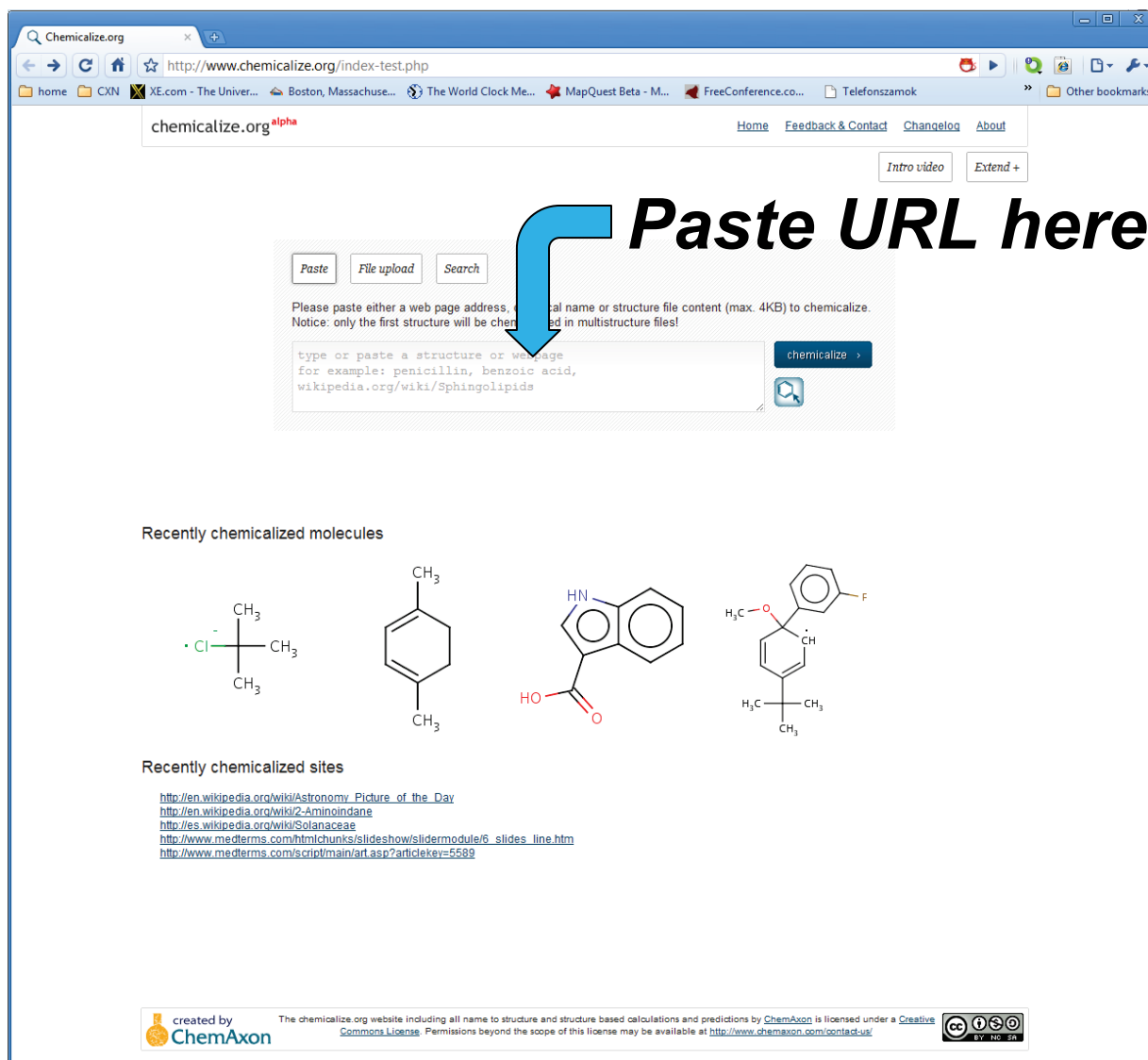
Free public service to:

- Add chemical structures to web pages in real time
- See predicted properties for discrete chemical structures
- Gather links of interest to chemists for post processing (search, analysis, reporting, fun...)
- Find chemical structures on the web (?)

chemicalize.org : What CXN is in the box?

1. *Name <> Structure*
 - *document extractor*
 - *name>structure generator*
2. *JChem + Marvin*
 - *Structure storage*
 - *Visualization*
3. *Calculator Plugins*
4. *JChem Webservices - search*

Add chemical structures to web pages



chemicalize.org^{alpha} Home Feedback & Contact Changelog About

Intro video Extend +

Paste File upload Search

Paste URL here

Please paste either a web page address, chemical name or structure file content (max. 4KB) to chemicalize.
Notice: only the first structure will be chemicalized in multistructure files!

type or paste a structure or message
for example: penicillin, benzoic acid,
wikipedia.org/wiki/Sphingolipids

chemicalize

Recently chemicalized molecules

CC(C)(C)Cl CC1=CC=C(C)C=C1 O=C1C=NC2=CC=CC=C12 CC(C)(C)C1=CC=C(C1)OC2=CC=CC=C2F

Recently chemicalized sites

http://en.wikipedia.org/wiki/Astronomy_Picture_of_the_Day
<http://en.wikipedia.org/wiki/2-Aminoindane>
<http://es.wikipedia.org/wiki/Solanaceae>
http://www.medterms.com/html/chunks/slideshow/slideshow/6_slides_line.htm
<http://www.medterms.com/scripts/main/art.asp?articlekey=5589>

created by ChemAxon

The chemicalize.org website including all name to structure and structure based calculations and predictions by ChemAxon is licensed under a Creative Commons License. Permissions beyond the scope of this license may be available at <http://www.chemaxon.com/contact-us/>

CC BY-NC-SA

Chemicalizing wikipedia List_of_organic_compounds” gives

W List of organic compoun...
http://www.chemicalize.org/?q=http://en.wikipedia.org/wiki/List_of_organic_compounds

The chemical names on this page were annotated by the [chemicalize.org](http://www.chemicalize.org) service. You can see the original page: http://en.wikipedia.org/wiki/List_of_organic_compounds.

WIKIPEDIA
The Free Encyclopedia

Article Discussion Read Edit View history Search

List of organic compounds

From Wikipedia, the free encyclopedia

This is a list of well-known **organic compounds**, including **organometallic** compounds, to stimulate the creation of Wikipedia articles. Note that purely **inorganic** compounds, minerals, and chemical elements are not included on this list. There are also no generic terms (e.g., carbohydrate) or mixtures of no fixed composition (e.g., naphtha, gasoline). Compounds and enzymes that are overwhelmingly of interest to biochemists, such as Cytochrome c peroxidase, are listed under [list of biomolecules](#).

For substances with a number prefix such as [2-Butanol](#) or [1,3-Cyclohexadiene](#), use the first **letter** of the name (in this case under B or C) to find the compound. Such names usually have the first **letter** capitalized in a title or at the beginning of a sentence.

Whilst most compounds are referred to by their IUPAC name, "traditional" names have also been kept where they are in wide use or of significant historical interest.

Contents: [A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#) [See also](#) [External links](#)

A

For substances with an A- or α - prefix such as [\$\alpha\$ -Terpinene](#), please see the parent page (in this case [Terpinene](#)).

- [Abietic acid](#) - C₂₀H₃₀O₂
- [Acenaphthene](#) - C₁₂H₁₀
- [Acenaphthoquinone](#)
- [Acenaphthylene](#) - C₁₂H₈
- [Acetaminide](#) - C₁₉H₂₂N₂O₂S
- [Acetaldehyde](#) - C₂H₄O
- [Acetamide](#)
- [Acetaminol](#)
- [Acetaminol](#)
- [Acetaminol](#)
- [Acetanilide](#)
- [Acetic acid](#)
- [Acetoguan](#)
- [Acetone](#)
- [Acetonitril](#)
- [Acetophen](#)
- [Acetylcho](#)
- [Acetylene](#)
- [N-Acetylgl](#)
- [Acetylsalicylic Acid](#) - C₉H₈O₄ (also known as Aspirin)
- [Acid fuchsin](#)
- [Acridine](#) — C₁₃H₉N
- [Acridine orange](#) - C₁₇H₁₉N₃
- [Acrolein](#) — C₃H₄O
- [Acrylamide](#) — C₃H₅NO
- [Acrylic acid](#) — C₃H₄O₂
- [Acrylonitrile](#) - C₃H₃N
- [Acryloyl chloride](#) - C₃H₃ClO
- [Acyclovir](#) - C₈H₁₁N₅O₃
- [Adamantane](#) - C₁₀H₁₆
- [Adenosine](#) - C₁₀H₁₃N₅O₄

Recognised chemical names are appended with dotted underline

Mousing over appended names pops the image

Clicking on image takes you to the "Data page" for that structure

Acpromazine

CN(C)CCCN1c2ccccc2Sc13ccccc13

Links are respected

W Acepromazine - Wikipedi... x

http://www.chemicalize.org/?url=http://en.wikipedia.org/wiki/Acepromazine

The chemical names on this page were annotated by the [chemicalize.org](http://www.chemicalize.org) service. You can see the original page: <http://en.wikipedia.org/wiki/Acepromazine>.

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article

Interaction
About Wikipedia
Community portal
Recent changes
Contact Wikipedia
Donate to Wikipedia
Help

Toolbox
Print/export

Languages
Deutsch
Français
Italiano
日本語
Norsk (bokmål)

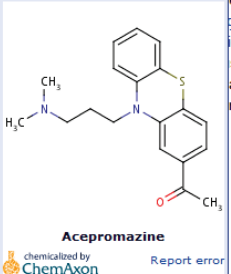
Article Discussion

Read Edit View history

Acepromazine

From Wikipedia, the free encyclopedia

Acepromazine or **acetylpromazine** (more commonly known as **ACP**, **Ace**, or by the trade names **Atravet** or **Acezine 2**, number depending on mg/l used in humans (the drug is frequently used in pharmaceutical practice in conjunction with and as such is not out of the preanes



Acepromazine
Report error
chemicalized by ChemAxon

1 Administration
1.1 Canine
1.1.1 Pot
1.2 Equine
1.2.1 Pre
1.3 Feline
2 References

Administration

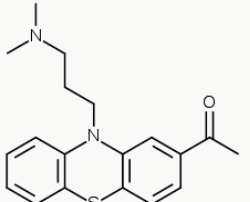
[edit]

Canine

[edit]

When used as a [premedication](#) it is commonly administered via the [subcutaneous](#) route.

Acepromazine



Systematic (IUPAC) name
1-[10-[3-(dimethylamino)propyl]-10H-phenothiazin-2-yl]ethanone

Identifiers

CAS number 61-00-7
ATC code N05AA04
PubChem CID 6077
DrugBank DB01614
ChemSpider 5852

Chemical data

Formula C₁₉H₂₂N₂O₅

*The URL database is populated by actual users visiting pages
“Human web crawling” technology*

Clicking on the structure image = Data page

chemicalize.org^{alpha} Open All Close All Manage calculations Layout: Custom acepromazine Search

Molecule

Name

IUPAC name: 1-[10-[3-(dimethylamino)propyl]-10H-phenothiazin-2-yl]ethan-1-one
Traditional name: Ace

Elemental Analysis

Formula: C₁₉H₂₂N₂O₂
Isotope formula: C₁₉H₂₂N₂O₂
Composition: C (69.9%), H (6.79%), N (8.58%), O (4.9%), S (9.82%)
Isotope composition: C (69.9%), H (6.79%), N (8.58%), O (4.9%), S (9.82%)
Mass: 326.456
Exact mass: 326.145284026

Major Microspecies

Major microspecies at pH=7.4:

pKa

Topology Analysis

Simple	Ring Counts	Path and distance
Atom count: 45 Bond count: 47 Cyclomatic number: 3 Chain atom count: 9 Chain bond count: 9 Asymmetric atom count: 0 Rotatable bond count: 5		

Polar Surface Area

Polar surface area: 23.55

logP

logP: 3.52

H-bond Donor/Acceptor

Geometry

Calculate Geometry

Lipinski-like filters

Lipinski's rule of five: yes
Bioavailability: yes
Ghose filter: yes
Lead likeness: yes
Muegge filter: yes
Veber filter: yes

Molecular Surface Area

Calculate Molecular Surface Area

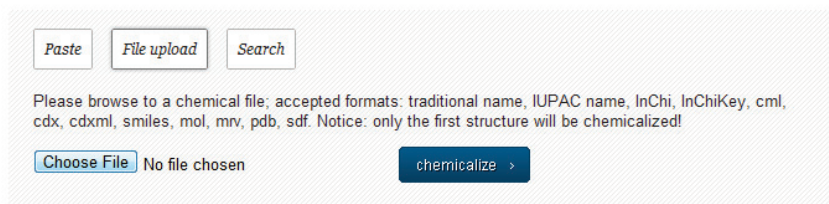
created by ChemAxon The chemicalize.org website including all name to structure and structure based calculations and predictions by ChemAxon is licensed under a Creative Commons License. Permissions beyond the scope of this license may be available at <http://www.chemaxon.com/contact-us/>

Customizable report for all ChemAxon calculations.

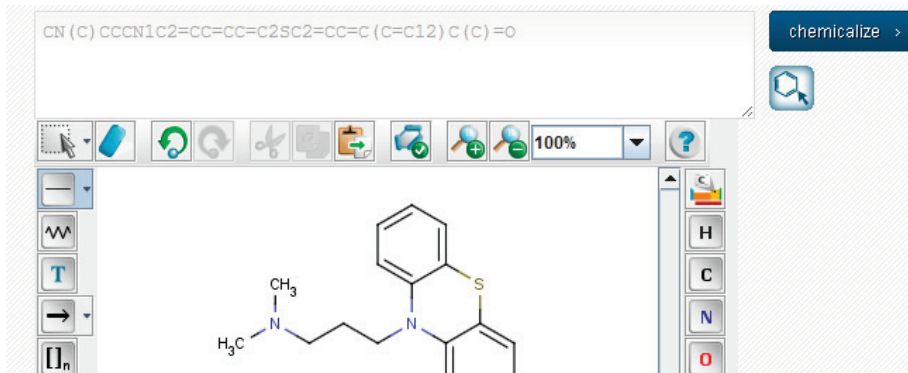
Users can move, open, close, expand calculation boxes and this is remembered on the next visit

Other ways to get to the Data page

1. Typing or pasting in your browser:
<http://www.chemicalize.org/?mol=acepromazine>
2. Typing “acepromazine” into the text box at [chemicalize.org](http://www.chemicalize.org)
3. Uploading a structure file (all formats) of acepromazine at [chemicalize.org](http://www.chemicalize.org)



4. Calling MarvinSketch at [chemicalize.org](http://www.chemicalize.org) and drawing/pasting acepromazine



What else about the Data page

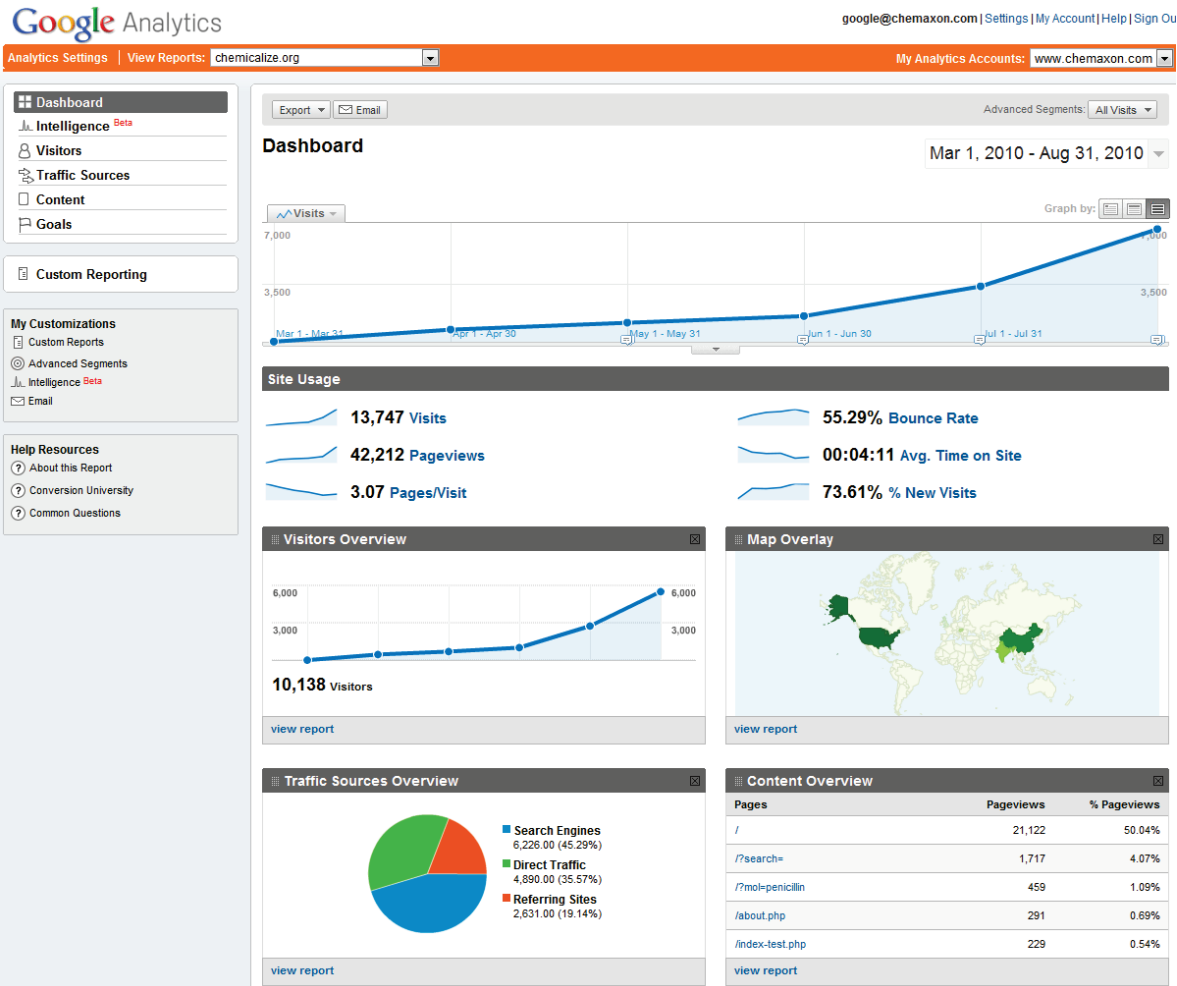
- Managing Calculation boxes, “Custom Layout”

The screenshot shows the top navigation bar of chemicalize.org with the following elements: 'Open All', 'Close All', 'Manage calculations' (dropdown), 'Layout: Custom' (dropdown), a search input field, and a 'Search' button. Below this, a grid of calculation boxes is displayed, categorized into: Basic, Protonation (Isoelectric Point +, logD +), Partitioning, Charge (Charge +, Polarizability +, Orbital Electronegativity +), Isomers (Tautomerization +, Stereoisomers +), Geometry, and Other (Huckel Analysis +, Refractivity +, Structural Frameworks +).

- Predefined Layouts.

The dropdown menu for 'Layout: Medicinal Chemist' is open, showing the following options: '- Select template -', 'Layout: Custom', 'Layout: Basic', 'Layout: Synthetic Chemist', 'Layout: Medicinal Chemist', and 'All calculations'.

Is it working...?



ChemSpider links out to data page (Sept 2010)



WordPress Plugin (Sept 2010):

In use at David Bradleys "Reactive Chemistry Blog".



Chris Swain releases Safari browser Plugin (Aug 2010)



Linguamatics I2E product links out to data page

What's next

- Search: sorting URL's relevance
- Structure images in more places (all URL's should have structures to represent them)
- Personalization "My chemicalize.org": Data page layouts, User History, parameterization of calculations...
- More browser and media engine plugins
- chemicalize.org landing page to be more representative of db content and interest of users "top structures, favorite frameworks, etc etc"
- Local server, private, licensed product
- HTML 5
- ***What else?***