



Europäisches
Patentamt
European
Patent Office
Office européen
des brevets

Possibilities for Searching with Chemical Structures in Conjunction with *esp@cenet*[®]



Nigel Clarke,
Vesna Hassler,
Marek Szubert

October 2007

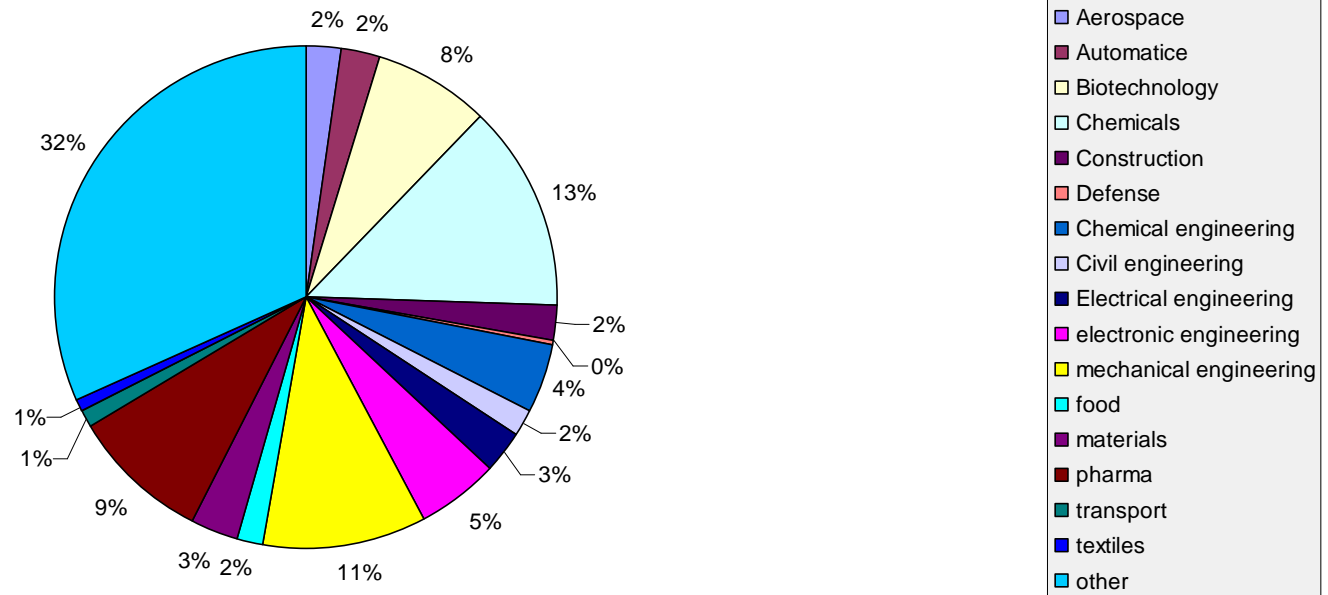


I wouldn't start from here if I were you



esp@cenet[®] User Profile; Industry Sector

What industry sector are you in?




esp@cenet[®]

- Chemical Structure Search
- Why?
- Large User Group (chemistry and pharma)



esp@cenet[®] advanced search



European Patent Office

Home | Contact

English | Deutsch | Français

esp@cenet

Help index ?

[Learn more about searching](#) [Get assistance](#) ↻

Quick Search

Advanced Search

Number Search

Last Results list

My patents list 3

Classification Search

Get assistance ↻

Quick Help

- » What does each database contain?
- » How many terms can I enter per field?
- » Can I search with a combination of words?
- » Can I use truncation or wildcards?
- » What are publication, application, priority and NPL reference numbers?
- » How do I enter publication, application, priority and NPL reference numbers?

Advanced Search

1. Database

Select patent database:

2. Search terms

Enter keywords in English

Keyword(s) in title:	<input type="text" value="plastic and bicycle"/>
Keyword(s) in title or abstract:	<input type="text" value="hair"/>
Publication number:	<input type="text" value="WO03075629"/>
Application number:	<input type="text" value="DE19971031696"/>
Priority number:	<input type="text" value="WO1995US15925"/>
Publication date:	<input type="text" value="yyyyymmdd"/>
Applicant(s):	<input type="text" value="Institut Pasteur"/>
Inventor(s):	<input type="text" value="Smith"/>
European Classification (ECLA):	<input type="text" value="F03G7/10"/>

esp@cenet[®] classification



European Patent Office



Home | Contact English Deutsch Français

Quick Search
Advanced Search
Number Search
Last Results list
My patents list **3**
Classification Search
Get assistance ↗

Quick Help

- » What is the ECLA classification system?
- » How do I enter a classification number?
- » Can I retrieve a classification using keywords?
- » Can I start a new search using the classifications listed?
- » Is the ECLA system updated?
- » How can I view the text of an ECLA class?

How do I use the Classification search? Get assistance ↗

Search the European classification

View Section **Find classification(s) for keywords** **Find description for a symbol**

Index **A B C D E F G H Y**

Next page: **A**

HUMAN NECESSITIES	A <input type="checkbox"/>
PERFORMING OPERATIONS; TRANSPORTING	B <input type="checkbox"/>
CHEMISTRY; METALLURGY	C <input type="checkbox"/>
TEXTILES; PAPER	D <input type="checkbox"/>
FIXED CONSTRUCTIONS	E <input type="checkbox"/>
MECHANICAL ENGINEERING; LIGHTING; HEATING; WEAPONS; BLASTING ENGINES OR PUMPS	F <input type="checkbox"/>
PHYSICS	G <input type="checkbox"/>
ELECTRICITY	H <input type="checkbox"/>
GENERAL TAGGING OF NEW TECHNOLOGICAL DEVELOPMENTS[N0403]	Y <input type="checkbox"/>

show notes Copy to searchform:

▲ top



esp@cenet[®] allowed inputs

- Words:
 - Keywords
 - Inventor names
 - Applicant names

esp@cenet[®] allowed inputs

- "Parameters"
 - Priority Number
 - Application Number
 - Publication Number

esp@cenet[®] allowed inputs

- Dates
 - Publication Date

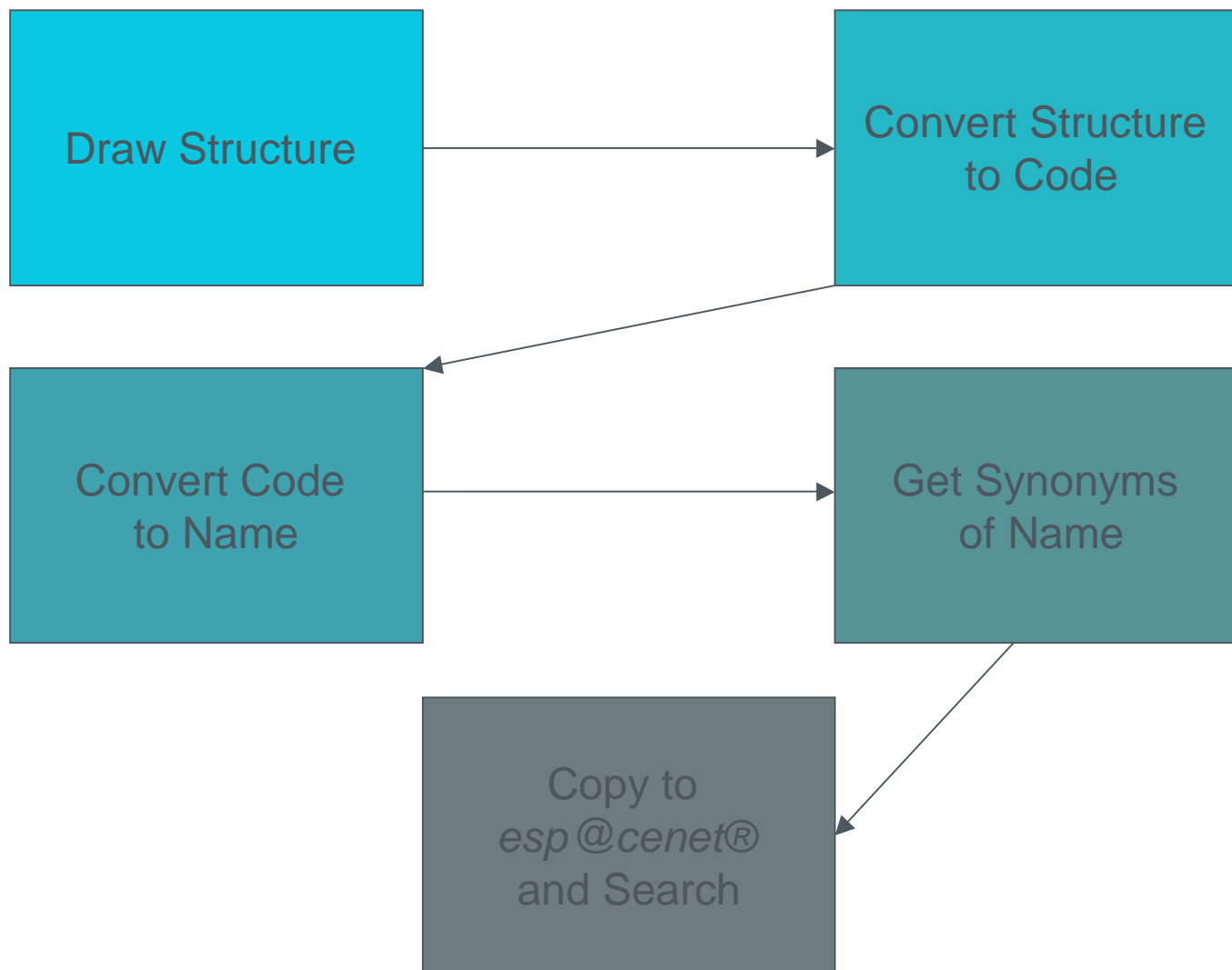
esp@cenet[®] allowed inputs

- Classifications
 - IPC
 - ECLA

esp@cenet[®] allowed inputs

- Only alphanumeric
- (wildcards allowed)

An Idea For Chemical Structure Searching



An Idea For Chemical Structure Searching

- Based on Public Domain Tools
 - Open Source
 - Freeware
 - Shareware
 - Proprietary



An Idea For Chemical Structure Searching

- Stand alone
- Does not alter the *esp@cenet*[®] application
 - c.f. translation engine
 - ECLA search

An Idea For Chemical Structure Searching

- Entry level
- To keep *esp@cenet*[®] attractive in the free sector
- Commercial suppliers can readily develop more sophisticated, added value tools



An open-source application

The screenshot displays the PubChem website interface. At the top left is the NCBI logo. The main header features the PubChem logo and the tagline "Information on biological activities of small molecules". A navigation bar includes links for HOME, SEARCH, SITE MAP, PubMed, Entrez, Structure, GenBank, PubChem, and Help. Below this is a "Compound Text Search" input field with a "GO" button. The "Structure Search" section is active, with radio buttons for "Basic" (selected) and "Advanced". It includes "Search", "Clear", and "Save Query" buttons. The "Search Input" section has radio buttons for "SMILES/SMARTS, InChI, CID or Formula" (selected), "Structure File", and "Saved Query". Below this is a "Sketch" button next to a drawing area. The "Search Type" section has a dropdown menu set to "Identical Structures". At the bottom left is a "Compound Subset" section with a plus icon and a help link. A mouse cursor is visible in the bottom right corner.

Draw structure using open source structure editor

The screenshot shows the PubChem Server Side Structure Editor V1.21 running in Microsoft Internet Explorer. The interface is divided into several sections:

- Top Bar:** Contains a dropdown menu for "Broadband" and a dropdown menu for "SMILES".
- Toolbar:** A row of icons for drawing and editing, including "New", "Undo", "Clear", "Delete", "Query", and various bond types (single, double, triple, aromatic, etc.).
- Ring and Functional Group Tools:** A row of icons for drawing rings (triangle, square, pentagon, hexagon, heptagon, octagon, benzene) and functional groups (nitro, carboxylic acid, ester, sulfonic acid).
- Periodic Table:** A grid of element symbols for selecting atoms. The Scandium (Sc), Yttrium (Y), and Lanthanum (Lu) columns are highlighted with dropdown menus.
- Export/Import Section:** Includes an "Export" button with a dropdown menu set to "MDL Molfile" and a "Done" button. Below it is a "Hydrogen" dropdown menu set to "Keep AsIs" and an "Import" button with a "Browse..." button next to it.

The status bar at the bottom of the browser window shows "Done" and "Internet".

Structure gives code - open source

PubChem Server Side Structure Editor V1.21 - Microsoft Internet Explorer

Broadband | SMILES | C1=CC=CC(=C1C(=O)O)C(=O)O

New Udo Cln Del Qry [Navigation icons]

[Bond types: single, double, triple, aromatic, etc.] [Ring types: triangle, square, pentagon, hexagon, heptagon, octagon, benzene]

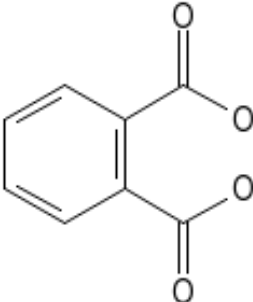
[Substituents: NO2, COOH, COOMe, SO3H]

H		?	?						He
Li	Be			B	C	N	O	F	Ne
Na	Mg			Al	Si	P	S	Cl	Ar
K	Ca	Sc	Sc	Ga	Ge	As	Se	Br	Kr
Rb	Sr	Y	Y	In	Sn	Sb	Te	I	Xe
Cs	Ba	Lu	Lu	Tl	Pb	Bi	Po	At	Rn

Export MDL Molfile Done

Hydrogen Keep AsIs

Import Browse...



Done Internet

Code Gives Chemical Name - open source

The screenshot displays the PubChem website interface. At the top left is the NCBI logo. The main header features the PubChem Compound logo. On the right, there are links for 'My NCBI', 'Sign In', and 'Register'. Below the header is a navigation bar with tabs for 'All Databases', 'PubMed', 'Nucleotide', 'Protein', 'Genome', 'Structure', 'PMC', 'PubChem', and 'Books'. A search bar is present with 'PubChem Compound' selected in the dropdown and 'for' in the input field. Below the search bar are buttons for 'Limits', 'Preview/Index', 'History', 'Clipboard', and 'Details'. The 'Display' section shows 'Summary' selected, 'Show 20', and 'Sort by'. Below this are filters for 'All: 1', 'BioAssay: 1', 'Protein3D: 1', and 'Rule of 5: 1'. The search results show '1: CID: 1017' with a link to 'Related Structures, Assays, Literature, Other Links'. The chemical structure of phthalic acid is shown, along with its name and properties: 'phthalic acid; o-phthalic acid; phthalate ...', 'IUPAC: phthalic acid', and 'MW: 166.13084 | MF: C8H6O4'. The left sidebar contains navigation links for 'About Entrez', 'Entrez Help', 'PubChem Help | FAQ', 'Courses', 'PubChem Substance Structures supplied by depositors', 'PubChem Compound Unique structures with computed properties', and 'PubChem BioAssay Bioactivity assay results supplied by depositors'.

Chemical Name Has Synonyms - open source

NCBI **PubChem** *Information on biological activities of small molecules*

HOME SEARCH SITE MAP PubMed Entrez Structure GenBank PubChem Help

Search PubChem Compound GO


Compound Summary:

O=C(O)c1ccccc1O



- CID: 1017** ⓘ ⓘ
- BioActivity: Summary** ⓘ
Inactive: 43 Links
Inconclusive: 1 Link
- Protein Structures: 2 Links** ⓘ
- NLM Toxicology:** ⓘ
Link1, Link2, Link3
- Substances:** ⓘ
All: 171 Links
Same: 29 Links
Mixture: 142 Links
- Related Compounds:** ⓘ
Same, Connectivity: 5 Links
- Similar Compounds: 253 Links** ⓘ
- Structure Search** ⓘ


MeSH Synonyms Properties Descriptors Category Exports

Get Synonyms - open source


 Depositor-Supplied Synonyms: (Total: 59) [?](#)

Display: [Next 10](#) | [All](#) | Sort:


- phthalic acid 
- o-phthalic acid
- phthalate 
- o-dicarboxybenzene
- 1,2-benzenedicarboxylic acid
- Acide phtalique
- o-benzenedicarboxylic acid
- Magnesium phthalate
- benzene-1,2-dicarboxylic acid
- Acide phtalique [French]

 Properties Computed from Structure: [?](#)

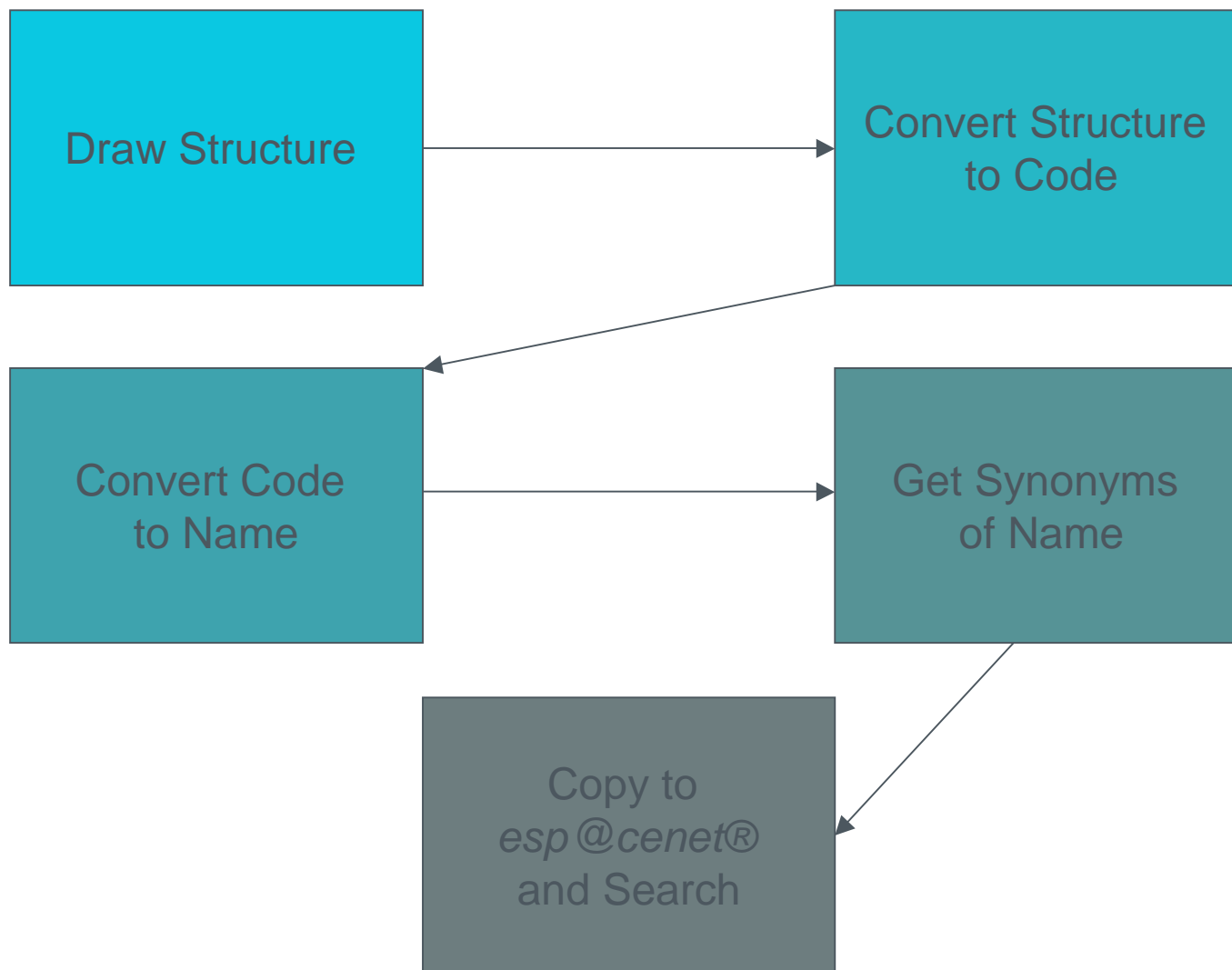
Molecular Weight: 166.13084 g/mol
Molecular Formula: C₈H₆O₄
Hydrogen Bond Donor Count: 2
Hydrogen Bond Acceptor Count: 4
Rotatable Bond Count: 2
Topological Polar Surface Area: 74.6
[+](#) more...

 Descriptors Computed from Structure: [?](#)

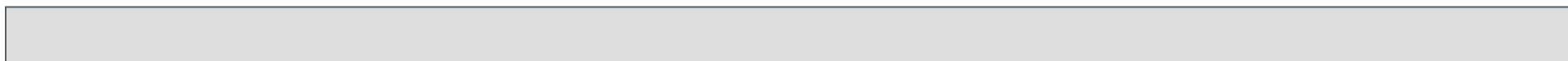
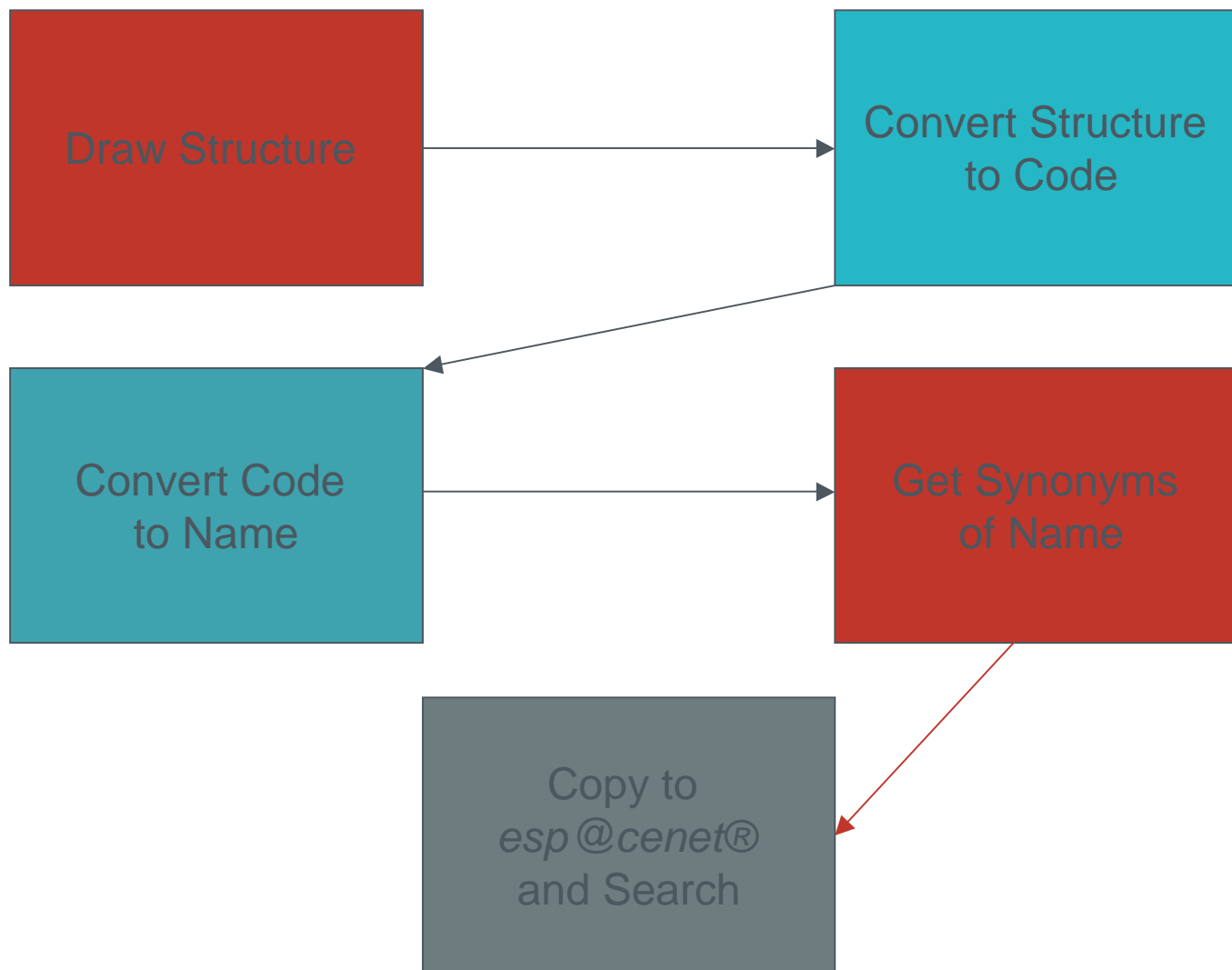
IUPAC Name: phthalic acid
Canonical SMILES: C1=CC=C(C(=C1)C(=O)O)C(=O)O
InChI: InChI=1/C8H6O4/c9-7(10)5-3-1-2-4-6(5)8(11)12/h1-4H,(H,9,10)(H,11,12)/f/h9,11H [?](#)

 Substance Category: [?](#)

An Idea For Chemical Structure Searching



An Idea For Chemical Structure Searching



I wouldn't start from here if I were you

