

# The National Chemical Database Service

Serin Dabb

Executive Editor, Data





# Royal Society of Chemistry

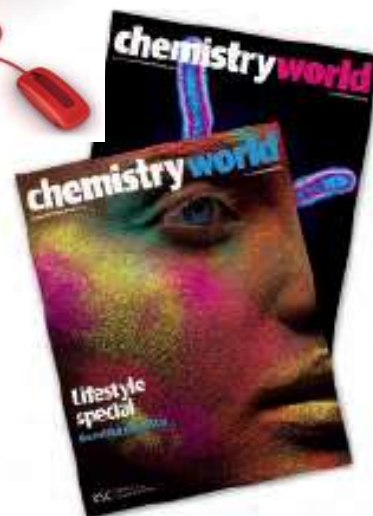
*World-leading chemistry community*

- Over 49,000 members
- Leading international not-for-profit publisher
- Scientific policy and education
- Conferences and events



# Leading international publisher

- Our publishing activities span books, e-books, journals, databases and magazines



# Who am I?

Databases

Literature  
Updating  
Services

THE  
**MERCK INDEX** *Online*

**MarinLit**

A database of the marine natural products literature



**SRU**




# Today's presentation

**CDS** National Chemical  
Database Service

[cds.rsc.org](https://cds.rsc.org)

**ChemSpider**  
Search and share chemistry

[chemspider.com](https://chemspider.com)



# What is the National Chemical Database Service (NCDS)?

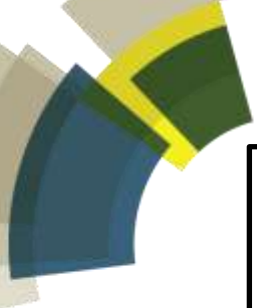
The NCDS is funded by the Engineering and Physical Sciences Research Council:



- 5 year contract (2013 - 2017)
- **Free to UK academic institutions**

## Vision

- A collection of **best-in-class databases and resources**
- An environment for **data validation, curation** and model building
- A platform for **open collaboration**
- An environment for **micropublishing**



## **UK Academia**

- Students: UG, PG
- Researchers, all levels
- Support staff

## **Chemistry Departments**

- Organic synthesis
- Inorganic chemistry
- Physical chemistry
  
- Crystallography
- NMR Spectroscopy

Who uses the  
NCDS?

## **Other departments**

**Material scientists / engineers**

**Biochemists / biologists**



# Access

- IP authentication (on campus)
- Shibboleth access (off campus)
- Unlimited users
- Some individual services require user names/password

# Currently NINE scientific resources

**CDS** National Chemical Database Service

In partnership with

**EPSRC**



[Home](#) [About CDS](#) [Learning Resources](#) [External Resources](#) [Help](#) [Contact Us](#)

The National Chemical Database Service brings together tools and resources for UK researchers in chemistry and related fields. All web-based services are freely accessible from any UK academic network. Further information about access to the NCDS can be found on the [About page](#).

NMR

Crystall-  
ography

Thermophysic  
al data

Crystall-  
ography

Commercial  
availability

Predicted  
physicochemi  
cal properties

ChemSpider

**ChemSpider**  
Search and share chemistry

An online database of molecules from  
>400 datasources (RSC).

[Further information](#)

Organic  
chemistry  
literature and  
reactions

Crystall-  
ography

[Introductory NCDS Video](#)



A brief introductory video to the  
National Chemical Database Service.

[Watch the video here](#)

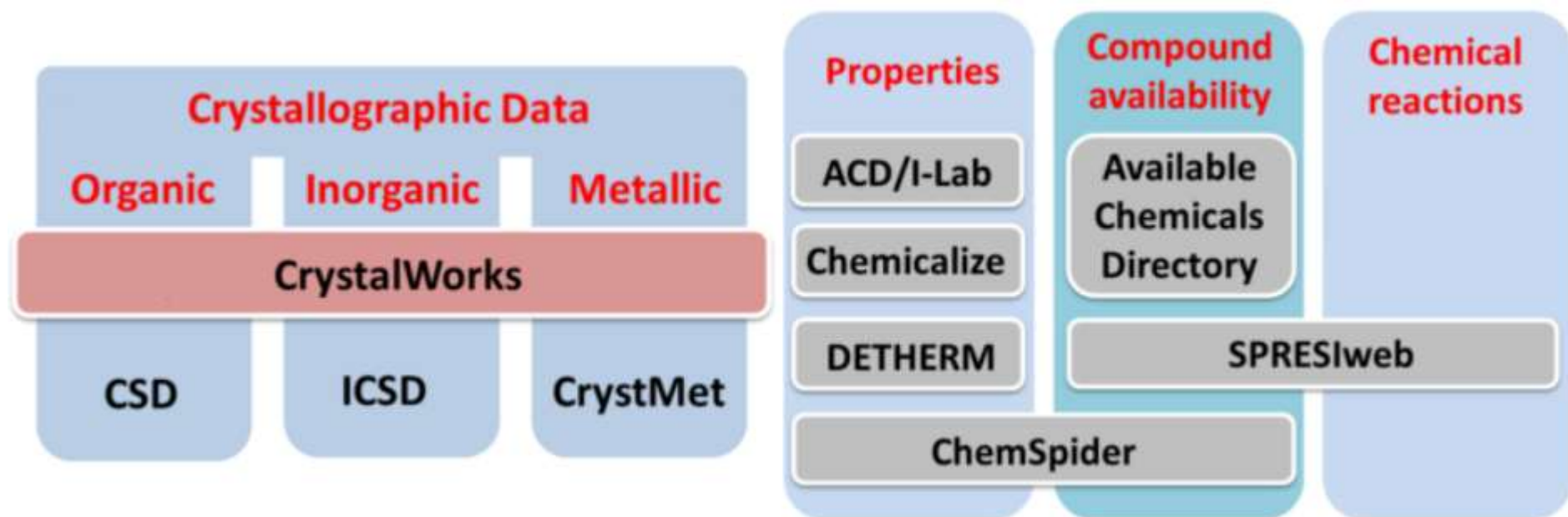
In partnership with the EPSRC

**EPSRC**

Engineering and Physical Sciences  
Research Council



# Which resource is right for a user?



# Crystallography: Organic and inorganic



## ICSD

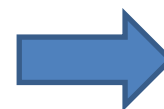
>Inorganic Crystal Structure Database

161,030 inorganic crystal structures

1,716 crystal structures of elements

Data from more than 1,900 periodicals

Records from 1913 – present



**CrystalWorks**



Sci-Tech Daresbury

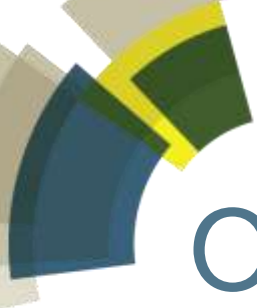


## CSD

>Cambridge Structural Database (from CCDC)

>600,000 small-molecule organic and organometallic crystal structures





# Organic chemistry: Reactions

## SPRESIweb



>Literature searching tool

>Search by name, trivial names, molecular formula, structure  
other data points

5.62 million molecules

4.34 million reactions

32 million factual data

Highlights commercially available molecules

Special keyword assignments



# Properties: Calculated and experimental

## **DETERM** thermophysical databases

- 4 M data sets,
- 129 K mixtures,
- 38 K pure compounds

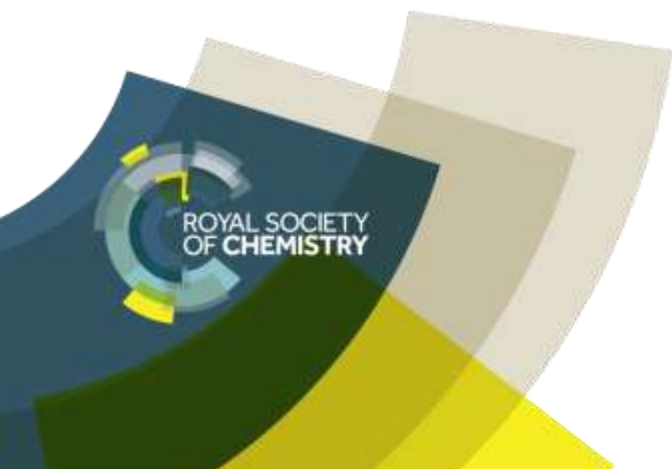


## **ACD/I-Lab**

- Structure searching, 160 K compounds
- Physicochemical properties
- NMR spectra prediction and database



# Other resources



The National Chemical Database Service is a gateway to chemical data for UK researchers in chemistry and related fields. All web-based services are freely accessible from any UK academic network. Further information about access to the NCDS can be found on the [About page](#).

## ACD/I-Lab



Physchem and NMR prediction and database (ACD/Labs Inc.).

[Further information](#)

## CSD



Organic and organometallic crystal structures (CCDC).

[Further information](#)

## DETERM



Database of thermophysical data for pure substances and mixtures.

[Further information](#)

## ICSD



>160,000 inorganic and related crystal structures (FIZ Karlsruhe GmbH).

[Further information](#)

## Available Chemicals Directory



Provides supplier information for building block molecules.

[Further information](#)

## Chemicalize



Physicochemical property prediction tools with Lipinski-like filters.

[Further information](#)

## ChemSpider



An online database of molecules from >400 datasources (RSC).

[Further information](#)

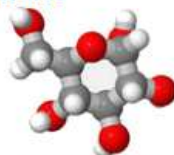
## SPRESIweb



Online chemical structure and reaction database (InfoChem GmbH).

[Further information](#)

## CrystalWorks



Crystallographic data from the CSD, ICSD and CrystMet (STFC Daresbury).

[Further information](#)

## Introductory NCDS Video



A brief introductory video to the National Chemical Database Service.

[Watch the video here](#)

In partnership with the EPSRC

**EPSRC**

Engineering and Physical Sciences  
Research Council



# NCDS: External resource list

We have compiled a list of **100** free online scientific resources **not provided** by the Royal Society of Chemistry

[cds.rsc.org/externalresources.asp](https://cds.rsc.org/externalresources.asp)

## Filter by Resource Type

- Biological chemistry
- Biological chemistry - nucleic acids
- Biological chemistry - proteins, enzymes
- Biological chemistry - small molecule
- Chemical availability
- Chemical semantics
- Chemical structure
- Computational chemistry
- General chemistry database
- Health and safety

- Kinetics, mechanisms
- NMR spectroscopy
- Materials science
- Organic systems
- Physical data
- Polymers
- Radiochemistry
- Solvents and solubility
- Spectroscopy
- Thermophysical properties
- Toxicology

## Filter by Publisher

- Academia
- EMBL-EBI
- EPA
- IUPAC
- NASA
- NIH
- NIST
- Other



# NCDS: External resource list

Show All

Sort by Resource Name (A-Z)

Sort by Publisher Name (A-Z)

[BindingDB](#) - Measured binding affinities. [Publisher: University of California, San Diego, USA]

[Biological Magnetic Resonance Data Bank](#) - NMR data of proteins, peptides, nucleic acids, and other biomolecules. [Publisher: University of Wisconsin, USA]

[Bordwell pKa Table](#) - [Publisher: University of Wisconsin, USA]

[BRENDA](#) - Enzyme data. [Publisher: Braunschweig University of Technology, Germany]

[ChEBI](#) - Binding, functional and ADMET bioactivity data, focused on small molecules. [Publisher: EMBL-EBI]

[ChemACX](#) - Compound supplier database [Publisher: PerkinElmer]

[ChEMBL](#) - Binding, functional and ADMET bioactivity data. [Publisher: EMBL-EBI]

[ChemDB](#) - Compound information [Publisher: University of California, Irvine, USA]

[ChemDB HIV](#) - Opportunistic Infection and Tuberculosis Therapeutics Database. This database contains data on the structure and activity of compounds that have been tested against HIV, HIV enzymes or opportunistic pathogens. [Publisher: National Institutes of Health (NIH), USA]

[ChemExper](#) - Compound supplier database [Publisher: ChemExper Inc., Belgium]

[Chemical Activity Predictor - GUSAR](#) - [Publisher: National Institutes of Health (NIH), USA]

[Chemical Components in the PDB](#) - [Publisher: EMBL-EBI]

[Chemical Effects in Biological Systems \(CEBS\)](#) - Data of interest to environmental health scientists. [Publisher: National Institutes of Health (NIH), USA]

[Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies](#) - [Publisher: National Aeronautics and Space Administration (NASA), USA]

[Chemical Structure Lookup Service](#) - Integrated searching of over 100 databases. [Publisher: National Institutes of Health (NIH), USA]

[ChemIDplus](#) - Information on 390,000 chemicals including synonyms and structures. [Publisher: National Institutes of Health (NIH), USA]

[ChemMine](#) - For analysing and clustering small molecules by structural similarities, physicochemical properties or custom data types. [Publisher: University of California, Riverside, USA]

[ChemSpider Synthetic Pages](#) - An interactive database of synthetic chemistry. [Publisher: Royal Society of Chemistry]

[ChemSynthesis](#) - Database of chemicals [Publisher: ChemSynthesis]

[CODATA](#) - Key values for thermodynamics [Publisher: International Council for Science : Committee on Data for Science and Technology]

[Computational Chemistry Comparison and Benchmark Database \(CCCBDB\)](#) - [Publisher: National Institute of Standards and Technology (NIST), USA]

[Constraint analysis](#) - Results from an analysis of constraints from 1834 NMR PDB entries. [Publisher: EMBL-EBI]

## Learning Resources

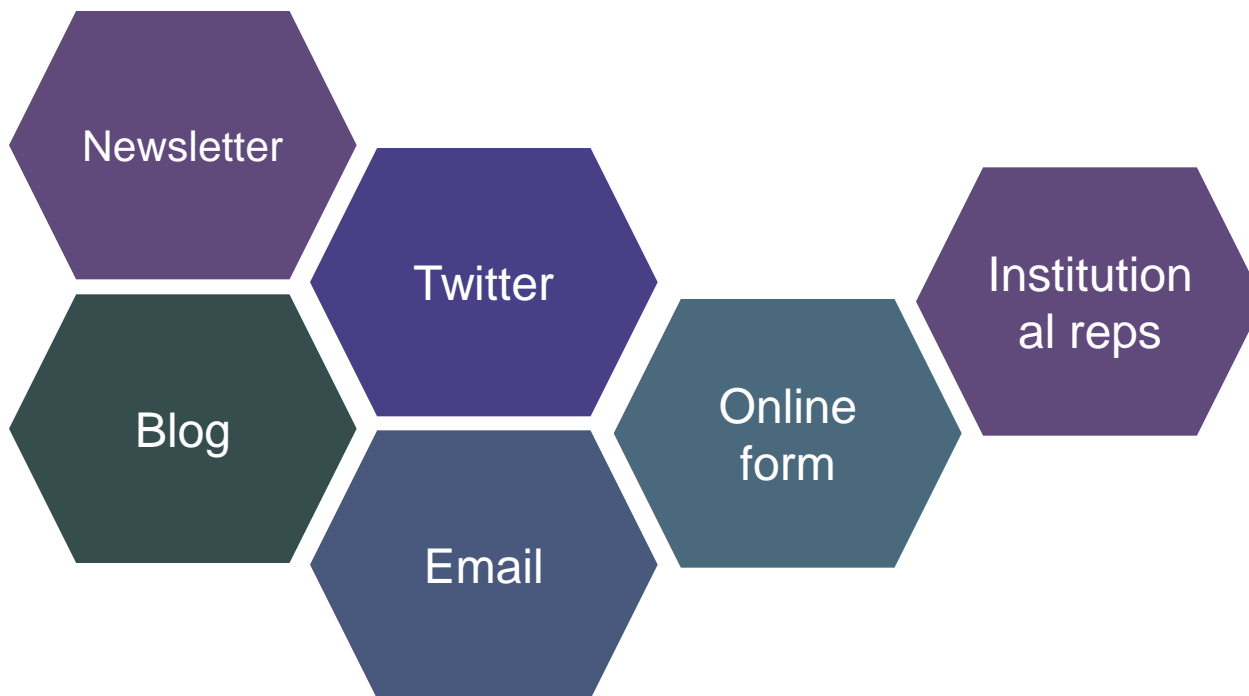
- Workshop material
- Flyers and fact sheets
- Videos
- Webinar series

The image shows three overlapping flyer images for different chemical databases. The top flyer is for CSD (Crystal Structure Database), the middle one is for Chemicalize (Chemical structure mine), and the bottom one is for DETHERM (Thermophysical Database). Each flyer provides a brief overview of the database's content and how to access it via the Chemical Database Service.





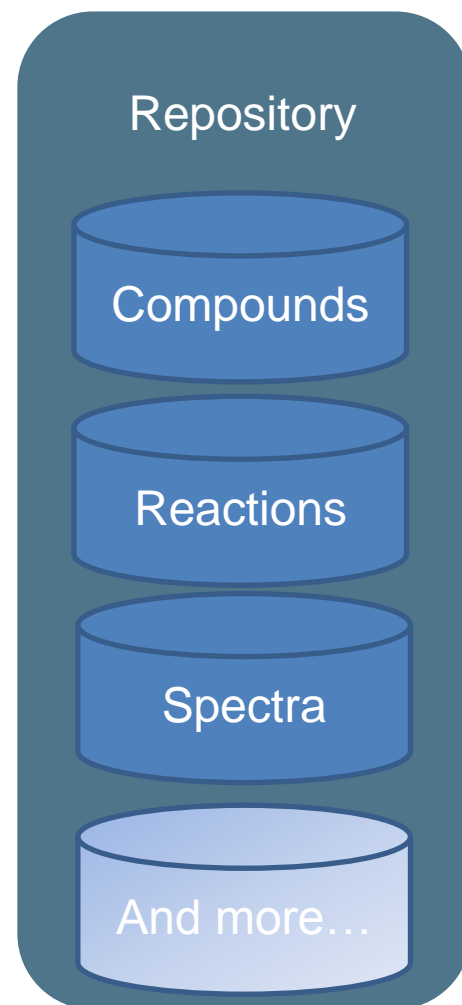
# Keeping in contact

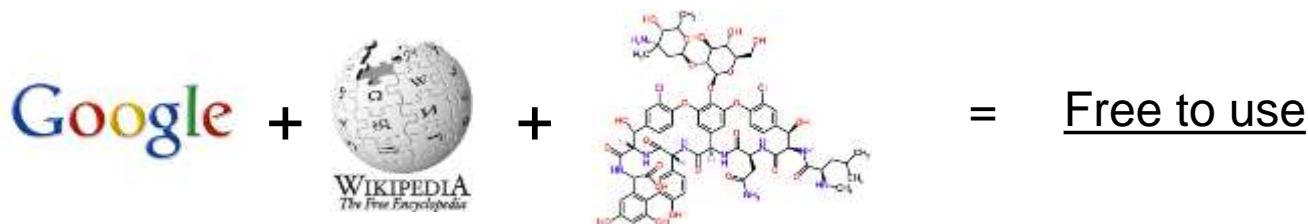




# Vision: Data repository

- Searchable, integrated, segregated repository of research data
- Chemically intelligent with appropriate metadata standards
- Data access including private, shared, embargoed and public





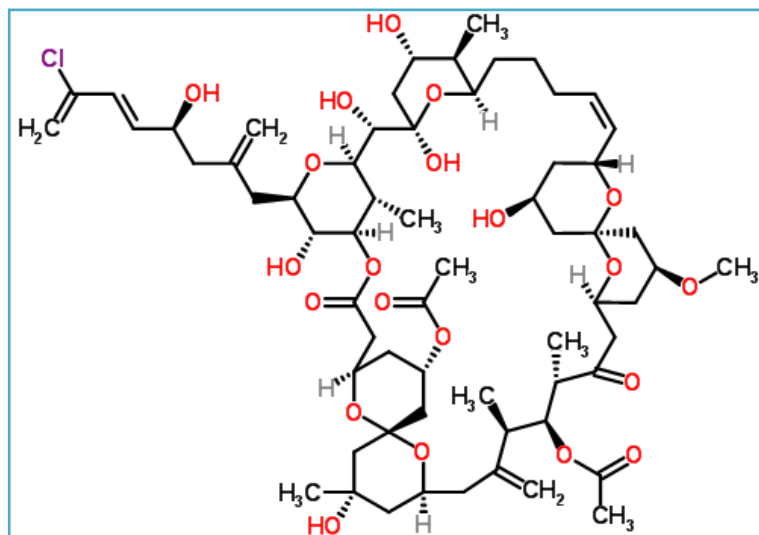
- 31 million structures
- Over 500 different data sources, including:
  - PubChem, ChEBI
  - GSK Malarial compounds
  - Chemical Suppliers Catalogues
  - Patents
  - RSC journals
- Physical properties, spectra, safety information and much more
- Access it anywhere there's internet

# What's in a name?

## Spongistatin 1


### Systematic Name:

(1S,3S,5R,9R,10R,11R,13R,14S,15R,17S,18R,19R,23Z,25R,27S,29S,31S,33S,36S,37S,38R,41S,43S,49S)-11-[(4S,5E)-7-chloro-4-hydroxy-2-methylideneocta-5,7-dien-1-yl]-10,14,15,17,27,43-hexahydroxy-31-methoxy-18,36,38,43,49-pentamethyl-39-methylidene-7,35-dioxo-8,12,45,46,47,48,50-heptaoxaheptacyclo[39.3.1.11,5.19,13.115,19.125,29.129,33] pentacont-23-ene-3,37-diyl diacetate



# New for 2015: Responsive design

[Home](#) [About us](#) [Web APIs](#) [Help](#) [Sign in](#)

 ROYAL SOCIETY OF CHEMISTRY


# ChemSpider


Search and share chemistry

[Simple](#) [Structure](#) [Advanced](#) [History](#)

## Search ChemSpider

Matches any text strings used to describe a molecule.



Systematic Name, Synonym, Trade Name, Registry Number, SMILES, InChI or CSID 

### What is ChemSpider?

ChemSpider is a free chemical structure database providing fast text and structure search access to over 34 million structures from hundreds of data sources.

### Search by chemical names

- Systematic names
- Synonyms
- Trade names
- Database identifiers

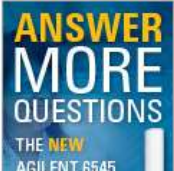
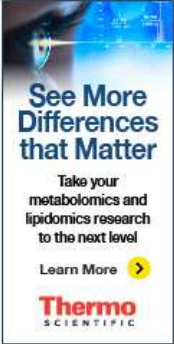
### Search by chemical structure

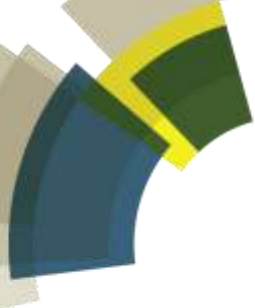
- Create structure-based queries
- Draw structures in the web page
- Use structure files from your computer

### Find important data

- Literature references
- Physical properties
- Interactive spectra
- Chemical suppliers

Advertisements

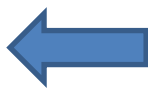




Basic chemical  
information



Community tagging



Much, much  
more  
information



# ChemSpider

Search and share chemistry

Simple Structure Advanced History

Found 1 result

Search term: **aspirin** (Found by approved synonym)

[COMMENT ON THIS RECORD](#)

Buttons:

## Aspirin

Molecular Formula	C <sub>9</sub> H <sub>8</sub> O <sub>4</sub>
Average mass	180.157 Da
Monoisotopic mass	180.042252 Da
ChemSpider ID	2157

Featured data source

The Merck Index Online has more data on this compound

More details:

[analgesic](#) [anti-inflammatory drug](#) [antipyretic](#) [antirheumatic drug](#) [+ TAG](#)

Names and identifiers

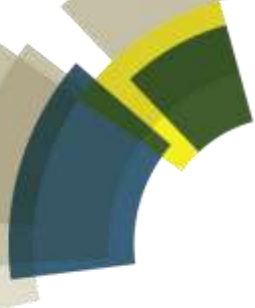
Names and Synonyms	Database ID(s)
<p>Validated by Experts, Validated by Users, Non-Validated, Removed by Users</p> <p><b>2-(Acetyloxy)benzoic acid</b></p> <p>200-064-1 <a href="#">[EINECS]</a></p> <p>2-Acetoxybenzenecarboxylic acid</p> <p>2-Acetoxybenzoesäure <i>[German]</i> <i>[ACD/IUPAC Name]</i></p> <p>2-Acetoxybenzoic acid <i>[ACD/IUPAC Name]</i></p>	

Properties Searches Spectra Vendors Articles More ▾

[EDIT](#)

[Give Feedback](#)

Generate Leads



Names and identifiers Properties Searches Spectra Vendors Articles More ▾

Names and identifiers Properties Searches Spectra Vendors Articles More ▾

Names and identifiers Properties Searches Spectra Vendors Articles More ▾

Names and identifiers Properties Searches Spectra Vendors Articles More ▾


Names and identifiers Properties Searches Spectra Vendors Articles More ▾


Names and identifiers Properties Searches Spectra Vendors Articles More ▾

Links & Reference RSC Journals RSC Books PubMed MeSH Literature

Results 1 - 10 of 102

<	1	2	3	...	10	11	>
---	---	---	---	-----	----	----	---

 Igor V. Tetko, Vsevolod Yu. Tanchuk, Tamara N. Kasheva, and Alessandro E. P. V. [Solubility of Chemical Compounds Using E-State Indices](#), *J. Chem. Inf. Comput. Sci.* **45**(12), 2615-2623  
[DOI: 10.1021/ci000392t]  
The molecular weight and electrotopological E-state indices were used to estimate aqueous solubility for a diverse set of 1291 organic compounds. The neural network provided highly predictive results with  $r^2 = 0.91$  and  $RMS = 0.62$ .

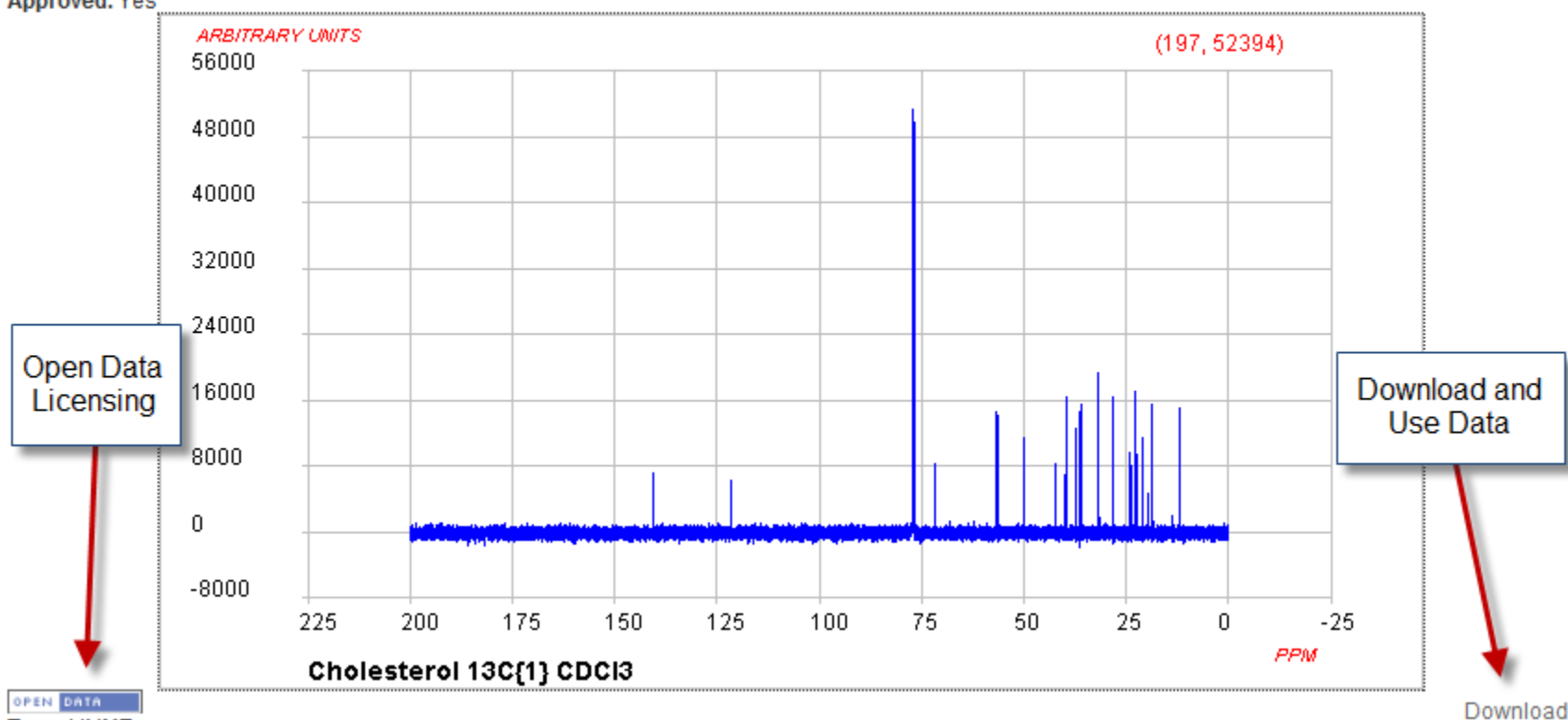
 Daniel F. Veber, Stephen R. Johnson, Hung-Yuan Cheng, Brian R. Smith, Keith W. Kopple. [Molecular Properties That Influence the Oral Bioavailability of Drug Candidates](#), *J. Med. Chem.* **45**(12), 2615-2623  
[DOI: 10.1021/jm020017n]  
275 drugs and their human oral bioavailability values have been given in the supplement. These values have in turn been taken from Goodman & Gilman's The Pharmacological Basis of Therapeutics, VIII and X Editions. The datasets are available from the QSAR World pages: <http://www.qsarworld.com/qsar-datasets-veber.php>

- Wikipedia
- Patents
- Crystal CIFs
- Media
- MeSH
- Pharma Links
- SimBioSys LASSO
- Images
- Data Sources
- Focused Libraries
- Curation



# Spectra

- Type: CNMR  
Approved: Yes





# Community curation

- Users can add
  - Descriptions, Syntheses and Commentaries
  - Links to PubMed articles
  - Links to articles via DOIs
  - Add spectral data
  - Add Crystallographic Information Files
  - Add photos
  - Add MP3 files
  - Add Videos

# ChemSpider

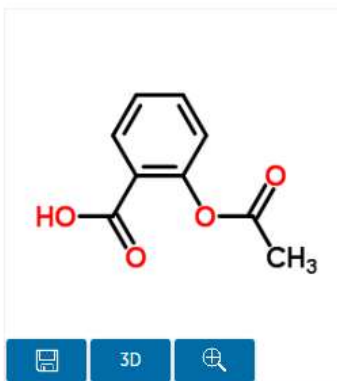
Search and share chemistry

Search ChemSpider

Simple Structure Advanced History

Found 1 result

Search term: **aspirin** (Found by approved synonym)



## Aspirin

Molecular Formula  $C_9H_8O_4$   
Average mass 180.157 Da  
Monoisotopic mass 180.042252 Da  
ChemSpider ID 2157



3D



### More details:

analgesic anti-inflammatory drug antipyretic antirheumatic drug

### Names and identifiers

Properties

Searches

Spectra

Vendors

Articles

More

Names and Synonyms

Database ID(s)

COMMENT ON  
THIS RECORD

Featured data source



The Merck  
Index Online  
has more data on  
this compound

Add data to this record

Identifier

Description

Image

Spectrum

Crystal CIF

Data source

DOI

PubMed ID

Advertisements



# ChemSpider SyntheticPages

[About](#)
[Search](#)
[Browse](#)
[Most Accessed](#)
[Recent Submissions](#)
[Help](#)
[My](#) 
[Submissions](#)

[Search](#)

Text search

Structure search

Search by Reaction name, Author name, or Compound

Search Reactions

Why use ChemSpider SyntheticPages?

- **It's free:** No licence or subscription required
- **Save time:** Find a reliable procedure that has already been optimised
- **Unique information:** Discover additional details direct from the researcher e.g. trouble-shooting tips, frequently encountered problems and scale up details

Why submit a procedure or method?

- **Quick and easy:** It's fast and simple to submit. Detailed instructions are [here](#).
- **Recognition:** All published procedures receive a DOI making them citable, discoverable and searchable.
- **Increase your visibility:** Submit procedures you have carried out, whether new or well-known, published or unpublished.

Example reactions

Biginelli reaction; N,N'-(4,4'-oxybis(4,1-phenylene))bis(6-methyl-4-phenyl-2-thioxo-1,2,3,4-tetrahydropyrimidine-5-carboxamide)

Sirin Gülden

9:1

24

[Advertise](#)
[Sponsor](#)

Advertisements

# Micropublishing

## Dehydration of 3,4-dihydro-5H-Benzo[cd]pyren-5-ol; 6H-Benzo[cd]pyrene

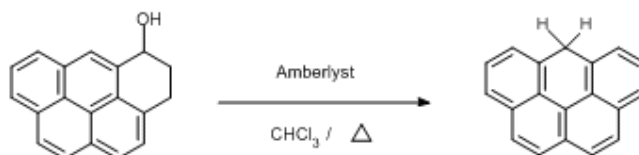
SyntheticPage 542

DOI: [10.1039/SP542](https://doi.org/10.1039/SP542)

Submitted Mar 15, 2012, published May 31, 2012

Anish Mistry ([a.mistry@warwick.ac.uk](mailto:a.mistry@warwick.ac.uk))

A contribution from Fox Group, Warwick University



### Chemicals Used

3,4-dihydro-5H-Benzo[cd]pyren-5-ol (prepared)

Amberlyst 15 (Sigma-Aldrich)

Chloroform

### Procedure

3,4-dihydro-5H-Benzo[cd]pyren-5-ol (0.1 g, 0.39 mmol) was dissolved in chloroform (30 ml) and Amberlyst 15 (0.1 g) added under a dinitrogen atmosphere. The reaction was heated to 30°C and left overnight under the inert atmosphere. The solution was then filtered to separate the Amberlyst and washed with chloroform. The combined solvents were removed under vacuum using a Rotary evaporator. The crude product was column chromatographed under a dinitrogen atmosphere eluting with 1:1 chloroform:petroleum ether 40-60°C. A white solid was obtained using this method (50 mg, 54%).



# Micropublishing

Deposit synthetic procedures on Synthetic Pages

- Share/Highlight your chemistry

Emphasis on reliable/robust chemistry

- Not so easy to gauge in traditional publications
- Platform allows users to comment/query protocol
- Share good practice (including safety information)

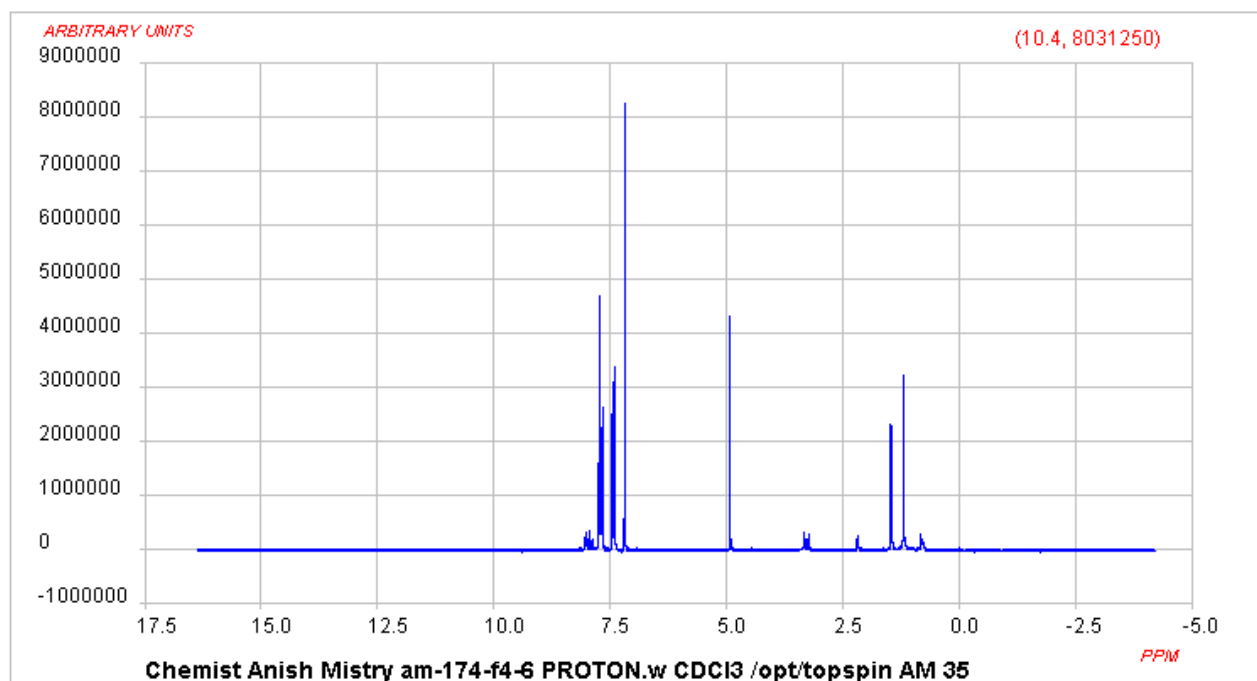
Records are **citable** (assigned DOIs)

# Interactive Data

## Supplementary Information

Proton NMR shows the symmetric dihydrogen species as the major product, the small peaks which can be observed in the spectrum (around 2, 3.5 ppm and aromatic region) are presumably other isomers of the compound.

[1H NMR spectrum of Olympicene - Click to view](#)




[1H NMR \(Jan11-2012.dx\)](#)


This page has been viewed approximately 1125 times since records began.

[Get structure file \(.cdx, .sk2, .mol\)](#)

# How to create a page

**ChemSpider**  
Synthetic pages

 ROYAL SOCIETY OF CHEMISTRY

About | Search | Browse | Most Accessed | Recent Submissions | Help | My  | Submissions | Ad

**Submit a Page**

**Source**

First name\*

Antony

Last name\*

Williams

Email

antony.williams@cher

Organization

Royal Society of Chem

Primary submitter

☒

Author 1

☐

Author 2

☐

Author 3

☐

Author 4

☐

Embargo Date ?

Group ?

**Abstracted submissions ?**





Abstracted submission (yes/no?) ☐

Original ID

Original URL


8+1


0

Advertise Sponsor

Advertisements

 APOLLO SCIENTIFIC

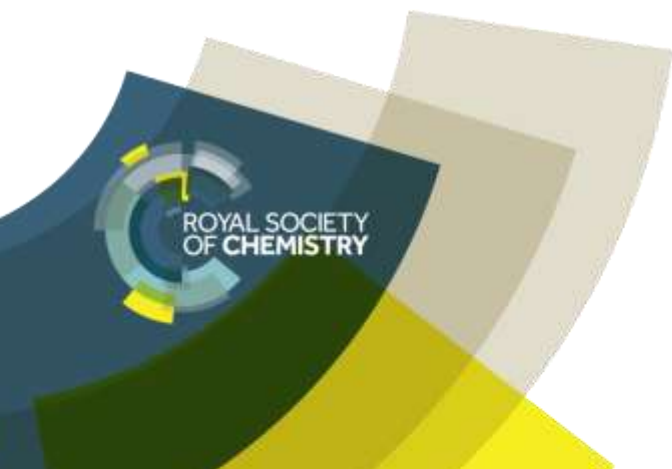
 ATTAIN  
SUPERIOR  
MASS SPEC  
PERFORMANCE  
GC/MS • LC/MS • ICP-MS  
LEARN MORE >  
Agilent

# Thank you

[cds@rsc.org](mailto:cds@rsc.org)

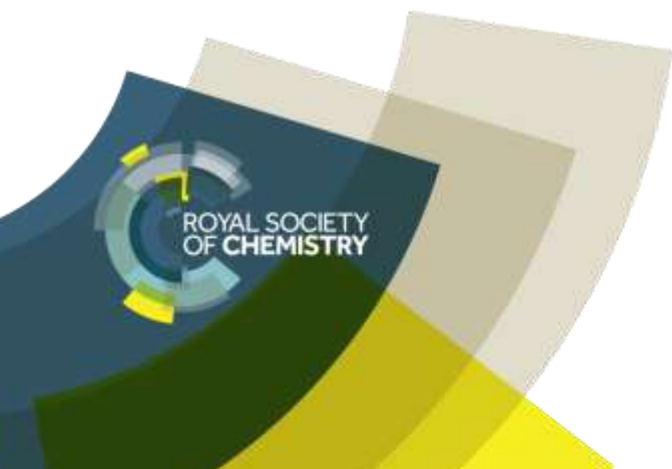
[ChemSpider@rsc.org](mailto:ChemSpider@rsc.org)

@SerinDabb



# Training at University of Bath

Linda Humphreys



## Course description

The RSC National Chemical Database Service provides access to a suite of specialist chemistry databases including the well-known crystallography databases CSD and ICSD. This session will give an overview of the service and look in detail at the non-crystallography databases: Spresiweb (topic searching); Detherm (thermophysical properties) and ACDiLab (physical and chemical properties and property prediction). This session will provide an overview of the databases and short demonstrations to help you get started with using the databases.

*What was most useful?*

- Overview of all the RSC tools
- All
- Finding out that the different databases exist
- ACD/I-Lab

*What was least useful?*

- Limited knowledge of crystallography databases
- Nothing

*Anything missing?*

- Bit more knowledge of crystallography databases

*What action will you take as a result of attending this workshop?*

- Search on RCS for crystal data
- Use the service