

NMRShiftDB – a spectral database

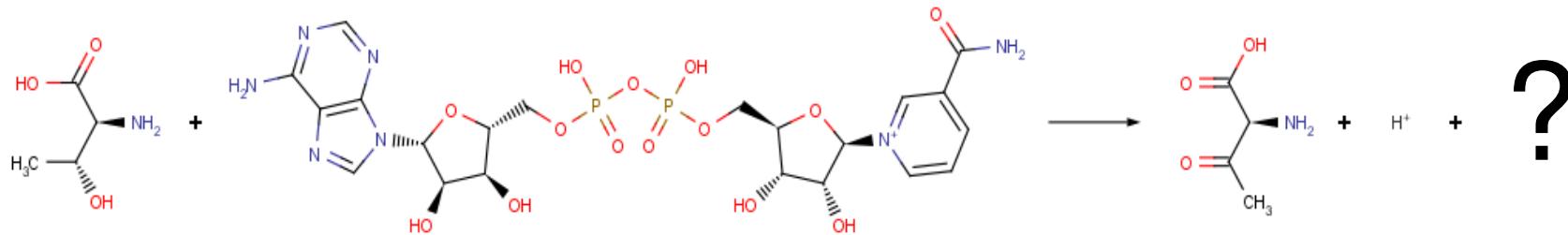
EBI Bioinformatics Roadshow

Paula de Matos



EBI is an Outstation of the European Molecular Biology Laboratory.

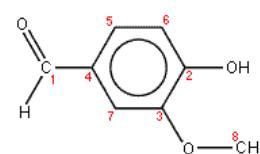
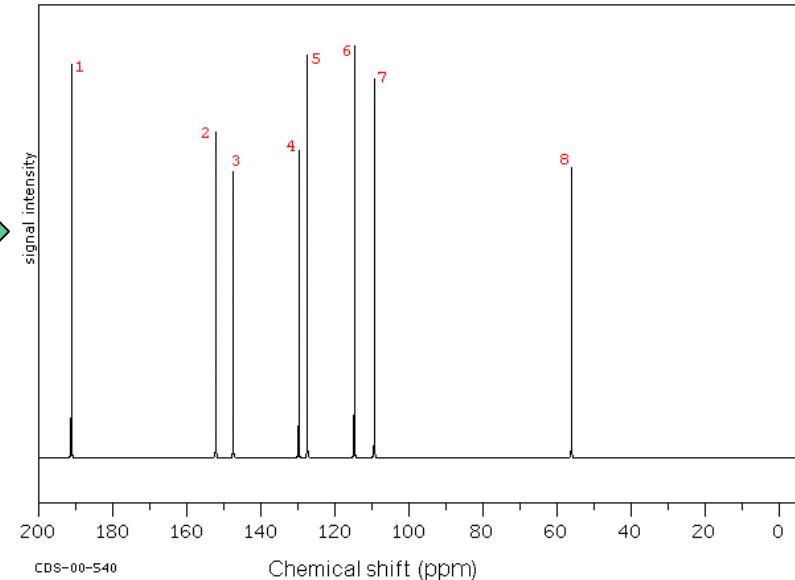
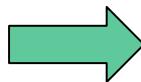
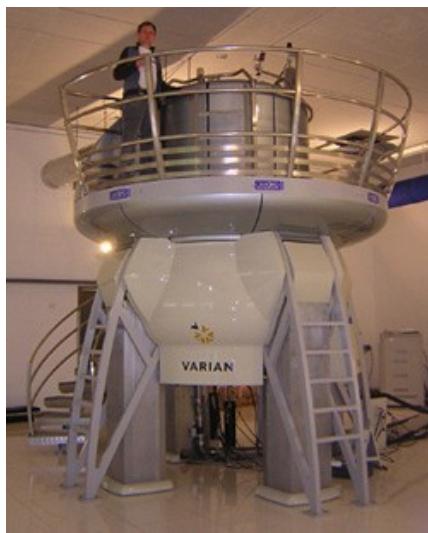
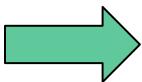
Identifying chemicals?



- Scientist needs to verify what a chemical is?
- Isolates a sample and then what?



NMR – Nuclear Magnetic Resonance



ppm	Int.	Location
191.21	955	1
152.18	791	2
147.50	692	3
129.77	746	4
127.49	975	5
114.75	1000	6 *
109.34	920	7 *
56.10	706	8

Introduction

- NMR - magnetic nuclei have in a magnetic field and applied electromagnetic (EM) pulse or pulses, which cause the nuclei to absorb energy from the EM pulse and radiate this energy back out.
- The energy radiated back out is at a specific resonance frequency which depends on the strength of the magnetic field and other factors.
- This allows the observation of specific quantum mechanical magnetic properties of an atomic nucleus.



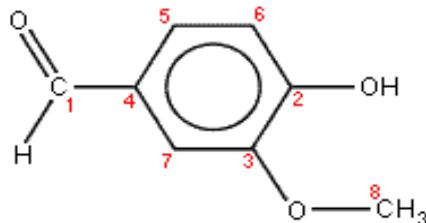
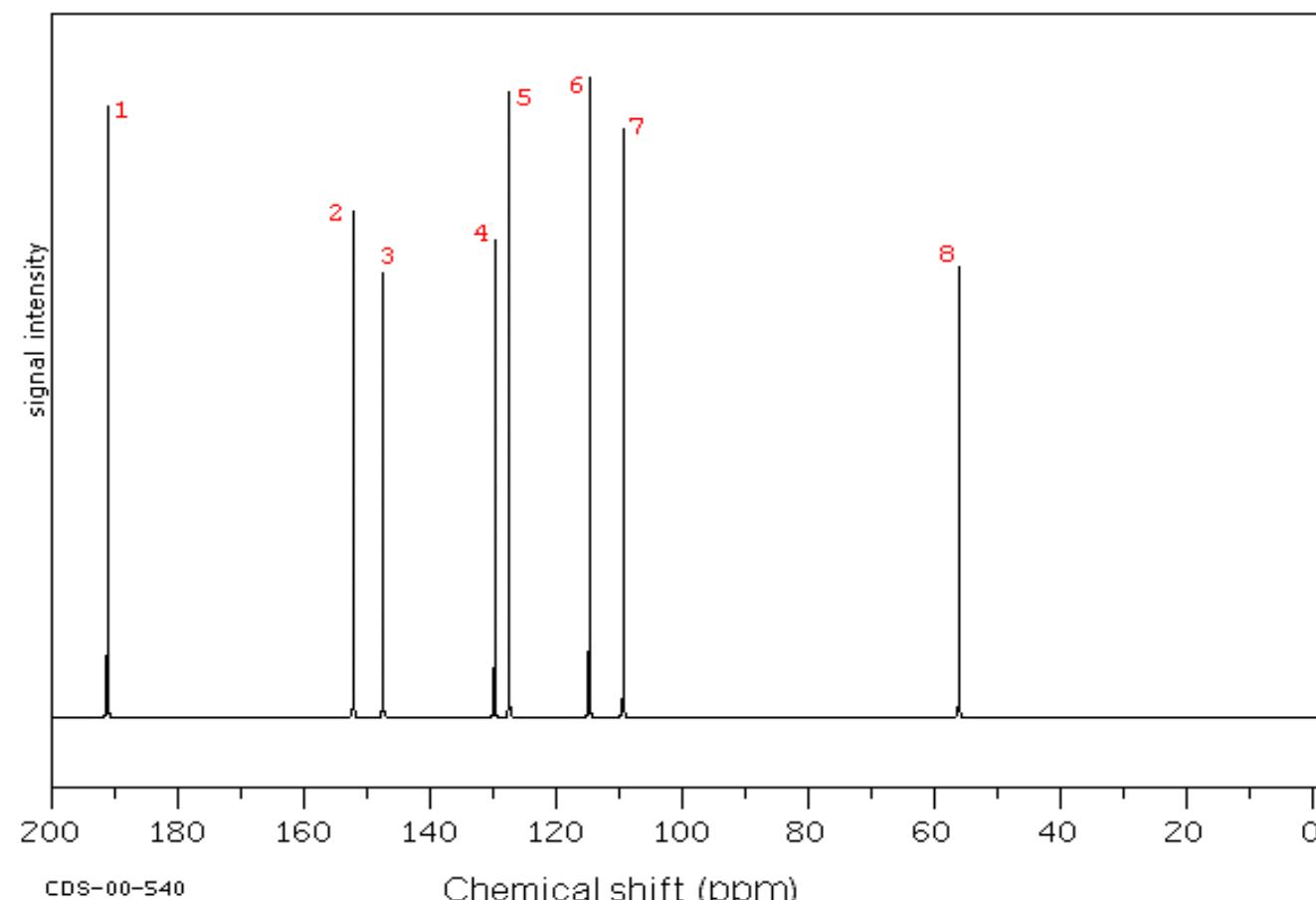
NMR vocabulary

- **Chemical Shift** – measured in ppm (parts per million). Measure of frequency relative to magnetic frequency

$$\delta = \frac{\text{difference in precession frequency between two nuclei}}{\text{operating frequency of the magnet}}$$

- Although frequency depends on the applied field the chemical shift is independent of it.
- **Intensity** – intensity of the signal at a particular frequency





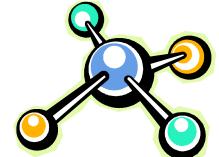
ppm	Int.	Location
191.21	955	1
152.18	791	2
147.50	692	3
129.77	746	4
127.49	975	5
114.75	1000	6 *
109.34	920	7 *
56.10	706	8

Favourite NMR Isotopes

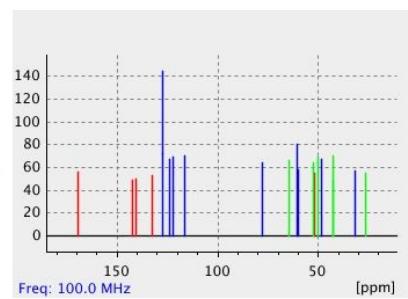
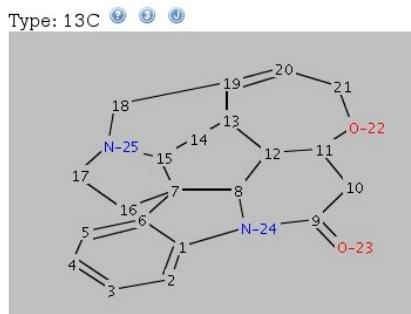
- Only some isotopes are detectable by NMR
- Hydrogen 1 (^1H) – natural abundance of hydrogen 1 is almost 100%
- Carbon 13 (^{13}C) - an important tool in chemical structure elucidation in organic chemistry.
- Carbon 13 has a natural abundance of 1.1% - not able to use Carbon 12 as its not detectable



NMRShiftDB – what is it about



- NMRShiftDB holds chemical structures and (assigned) NMR peak spectra.
- NMR peak spectra can be used as fingerprints of compounds.
- Therefore, they can be used for structure elucidation and verification.



Atom No.	Mult.(coupling const.)	Meas. Shift	Intensity	expt-0
1	S	140.4	0.616352	141.99
2	D	116.3	0.877358	115.96
3	D	127.4	0.902516	128.33
4	D	124.2	0.839623	123.98
5	D	122.2	0.867925	122.03
6	S	132.7	0.666667	132.39
7	S	52.0	0.691824	51.73
8	D	60.6	1.0	59.84





General properties

- Available via the web:
<http://www.ebi.ac.uk/nmrshiftdb> or
<http://www.nmrshiftdb.org>
- Community built
- All data are freely available under the GNU FDL and bulk downloadable for use in your own research
- This tutorial will cover the use of the web interface

NMRshiftDB home page

**NMRShiftDB****Current usage of NMRshiftDB is:**

Registered Users: 2083

Structures which can be searched: 34824

Spectra: Measured 38747, calculated 550

Username: [Login](#)Password: [Logon with SSL](#) Use cookies for persistant login[Create New Account](#)

(Only necessary for contributing data)

[Forgot password?](#)**Problems using NMRShiftDB? See our [tips on browsers to use!](#)**[Home](#) [Search](#) [Results](#) [Predict](#) [Submit](#) [Review](#) [Wishlist-Guestbook](#) [Help](#)

NMRShiftDB Links

[Developers' page](#)[Sponsoring](#)[Media coverage](#)[Static name list](#)[Links](#)[FAQ](#)[Contact](#)

Hall of Fame

Name	Contributions
1 E. Willighagen	1094
2 S. Dathe	505
3 P. Braeutigam	439
4 S. Kuhn	393
5 N. Prakash	350
6 B. Patel	305
7 M. Gericke	181
8 N. Kuznik	120
9 K. Bohn	111
10 R. Ellinger	76
11 A. Dransfeld	56
12 K. Bartusseck	26
13 M. Mitchell	20

About NMRShiftDB

NMRShiftDB is a NMR database (web database) for organic structures and their nuclear magnetic resonance (nmr) spectra. It allows for spectrum prediction (¹³C, ¹H and other nuclei) as well as for searching spectra, structures and other properties. Last not least, it features peer-reviewed submission of datasets by its users. The NMRShiftDB software is open source, the data is published under an open content license. Please consult the [documentation](#) for more detailed information.

News about NMRShiftDB

Standalone client released 2009-09-07 13:59 - [NMRShiftDB](#)

The 1.0 release of our standalone client called Specclipse is available for download from [here](#). It offers offline editing of data and download and submit facilities. It is based on Bioclipse 2.0.

[Read More »](#)**Server Consolidation** 2009-07-08 11:41 - [NMRShiftDB](#)

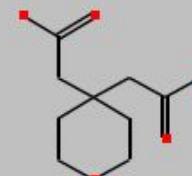
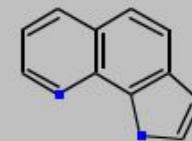
We have consolidated the NMRShiftDB servers to run at European Bioinformatics Institute. The URL [here](#) and all existing server URLs redirect to the server, so users should not notice that except from the URL in the browser.

[Read More »](#)**Bugfix release** 2009-04-07 10:42 - [NMRShiftDB](#)

Bug fix release 1.3.4 is now available. It fixes some bugs and cleans up ID handling. See [here](#) for details.

[Read More »](#)

Latest Additions



Hands On Tutorial!

EMBL-EBI



EBI is an Outstation of the European Molecular Biology Laboratory.

Example 1: Searching by spectrum

NMRShiftDB Current usage of NMRShiftDB is:

Registered Users: 2000
Structures: 34824
Spectra: 550

Search

Problems using NMRShiftdb? See our [tips on browsers to use !](#)

Username: Login
Password:
 Use cookies for persistant login
[Logon with SSL](#)

Home | **Search** | Results | Predict | Submit | Review | Wishlist-Guestbook | Help

Search by Spectrum **Search History**

Switch to expert search mode

Browse all structures

Input list :
155.3
151.6
147.5
107.8
144.3
27.8
29.6
33.5

Enter spectrum data

Repeat Type of search Mode Value Hits
 not and or
Repeat marked search(es) [?](#)

Clear history

Subspectrum Complete

Search by spectrum

Subspectrum or complete

Or choose jcamp file:
 [Browse...](#) [Upload file](#)

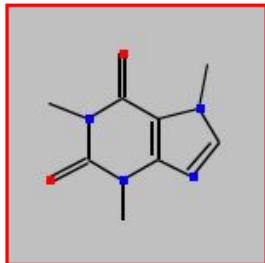


Search Results

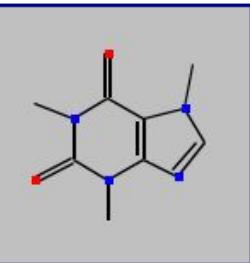
There were more than 300 results. The rest was cut off ! ⓘ

Type of search	Mode	Value
total similarity spectrum search	--	27.8 29.6 33.5 107.8...

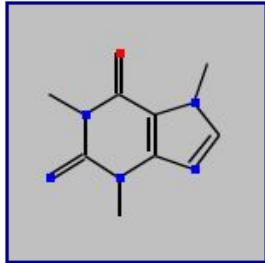
Results: 300

Browse: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) >>**Next structure**

Similarity: 100.00 %



Similarity: 95.48 %



Similarity: 85.80 %



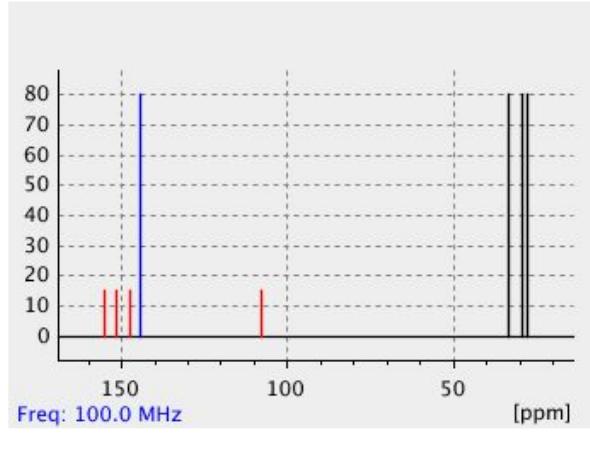
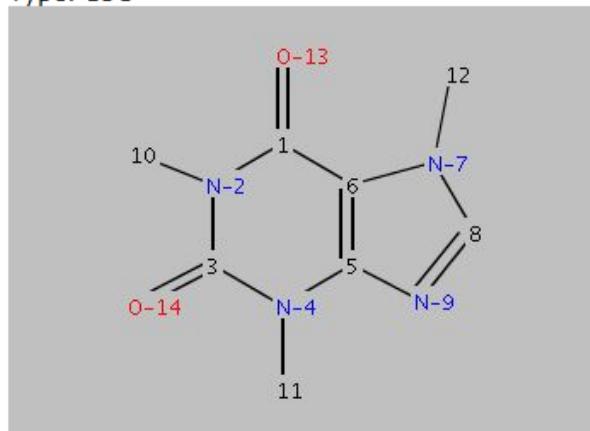
Similarity: 82.49 %

**Bookmarks**

You could bookmark structures if you were logged in!

Details**Spectral Data** [Additional Data](#)

Search for complete spectrum: Similarity measure for the complete spectrum in this record is 100.0.

Type: ¹³C ⓘ ⓘ ⓘ

Atom No.	Mult.(coupling const.)	Meas. Shift	Input Shift	Diff. M-I	expt-0 sdb
1	S	155.3	155.3	0.00	155.32
3	S	151.6	151.6	0.00	148.67
5	S	147.5	147.5	0.00	151.66
6	S	107.8	107.8	0.00	107.51
8	D	144.3	144.3	0.00	141.57
10	Q	27.8	27.8	0.00	29.7
11	Q	29.6	29.6	0.00	27.88
12	Q	33.5	33.5	0.00	33.57

Threshold is 12.50

Details

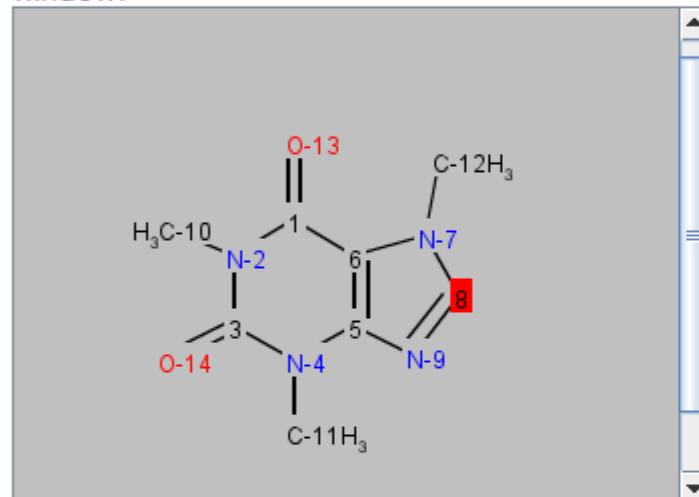
User input (2D) ▾

Show with this coordinate set

Spectral Data Additional Data

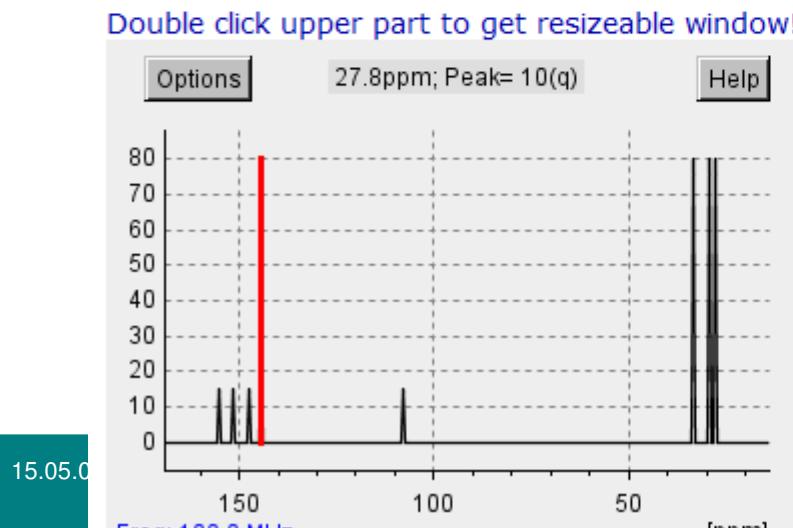
Search for complete spectrum: Similarity measure for the complete spectrum in this record is 100.0.

Type: 13C ⓘ Double click to get resizable window!



Atom No.	Mult.(coupling const.)	Meas. Shift	Input Shift	Diff. M-I	expt-0 sdbs
1	S	155.3	155.3	0.00	155.32
3	S	151.6	151.6	0.00	148.67
5	S	147.5	147.5	0.00	151.66
6	S	107.8	107.8	0.00	107.51
8	D	144.3	144.3	0.00	141.57
10	Q	27.8	27.8	0.00	29.7
11	Q	29.6	29.6	0.00	27.88
12	Q	33.5	33.5	0.00	33.57

Threshold is 12.50



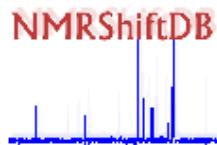
Details

Spectral Data Additional Data

Molecule	10016316
Chemical name(s)	Coffein
Chemical formula	C ₈ H ₁₀ N ₄ O ₂
Molecular weight	194.191
Number of double bond equivalents (DBEs)	6.0
Number of all rings, size of smallest set of smallest rings	3, 2
Canonical name(s)	<ul style="list-style-type: none"> • O=C2c1c(ncn1C)N(C(=O)N2C)C (SMILES) • 1,3,7-trimethyl-3,7-dihydro-1H... (truncated) (IUPAC from ACD/Name) • 1H-purine-2,6-dione, 3,7-dihyd... (truncated) (Index from ACD/Name) • InChI=1/C8H10N4O2/c1-10-4-9-6-... (truncated) (INChI) • RYYVLZVUVIJVGH-UHFFFAOYAW (InChI Key)
CAS-Number	58-08-2
Additional information	Deposition in PubChem ; Compound in PubChem ; ChEBI ;
Spectrum	30127729 Rating: 10
Type	13C
Measurement conditions	
Assignment Method	Unreported
Field Strength [MHz]	Unreported
Temperature [K]	Unreported
Solvent	Unreported
Literature	S. Berger; S. Braun; H.-O. Kalinowski: 13-C-NMR-Spektroskopie, New York: Thieme Verlag 1984.
Additional comments	624.MOL; / * changed /
Additional information	
Spectrum categories	

Additional
molecule data

Example 2: Search by JCAMP file



Current usage of NMRShiftDB is:

Registered Users: 2083

Structures which can be searched: 34824

Spectra: Measured 38747, calculated 550

Username: Login

Password: [Logon with SSL](#)

Use cookies for persistant login

Problems using NMRShiftdb? See our [tips on browsers to use !](#)

[Home](#) [Search](#) [Results](#) [Predict](#) [Submit](#) [Review](#) [Wishlist-Guestbook](#) [Help](#)

Search by Spectrum

Search History

[Switch to expert search mode](#)

Repeat	Type of search	Mode
<input type="checkbox"/>	total similarity spectrum search --	26.8 3
<input type="checkbox"/>	total similarity spectrum search --	27.8 29

[Browse all structures](#)

Input list [?](#):

133.28;0.14
123.85;0.12
115.17;0.13
113.56;0.16
70.45;0.13
46.55;0.16
40.15;0.15
39.93;0.44
39.73;0.94
39.51;1.0
39.31;0.94

Subspectrum
 Complete

[Search by spectrum](#)

not and or

[Repeat marked search\(es\)](#) [?](#)

[Clear history](#)

Upload
JCAMP file

Or choose jcamp file:

C:\Users\Paula\Documents

[Browse...](#)

[Upload file](#)

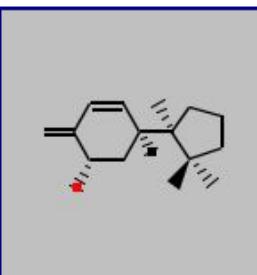
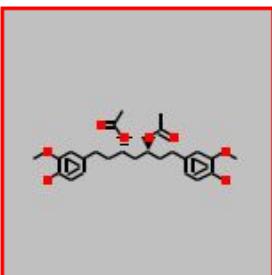


Search Results

There were more than 300 results. The rest was cut off ! 

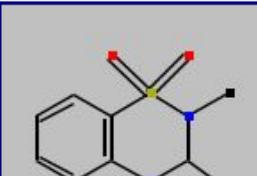
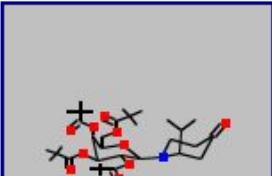
Type of search	Mode	Value
total similarity spectrum search	--	21.26;0.14 30.58;0.1...

Results: 300

Browse: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) >>**Next structure**

Similarity: 49.01 %

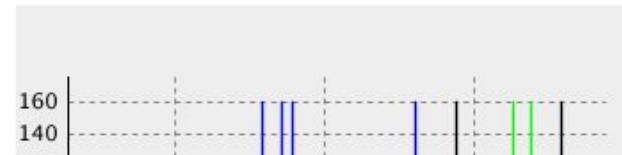
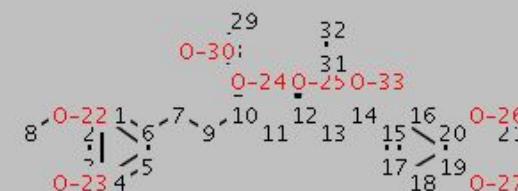
Similarity: 44.82 %

**Bookmarks**

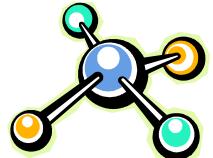
You could bookmark structures if you were logged in!

Details**Spectral Data** [Additional Data](#)

Search for complete spectrum: Similarity measure for the complete spectrum in this record is 49.01.

Type: ¹³C   

Atom No.	Mult.(coupling const.)	Meas. Shift	Input Shift	D
1	D	110.9	113.56	2
2	S	146.3		0
3	S	143.7		0
4	D	114.2	115.17	0
5	D	120.8	123.85	3
6	S	133.2	133.28	0
7	T	31.2	30.58	0
8	Q	55.8		0
9	T	36.7	38.89	2
10	D	69.7	70.45	0
11	T	38.6	39.11	0
12	D	69.7	70.45	0
13	T	36.7	38.89	2
14	T	31.2	30.58	0
15	S	133.2	133.28	0
16	D	110.9	113.56	2
17	D	120.8	123.85	3
18	D	114.2	115.17	0
19	S	143.7		0
20	S	146.3		0



Example 3: Structure search

- Given a chemical structure, search for similar structures

Search by Structure

Draw structure

Get mol file

Structure file

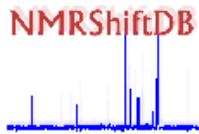
Browse... Import file

Import text by copy & paste

Structure search

Substructure Search
Similarity Search
Identity Search ?
Clear
Import from structures history
Search by structure

Identity search results



Current usage of NMRShiftDB is:

Registered Users: 2083

Structures which can be searched: 34824

Spectra: Measured 38747, calculated 550

Username:

[Login](#)

Password:

[Logon with SSL](#)

Use cookies for persistant login

[Create New Account](#)

(Only necessary for contributing data)

[Forgot password?](#)

Problems using NMRShiftdb? See our [tips on browsers to use !](#)

[Home](#) [Search](#) [Results](#) [Predict](#) [Submit](#) [Review](#) [Wishlist-Guestbook](#) [Help](#)

Search Results



Bookmarks

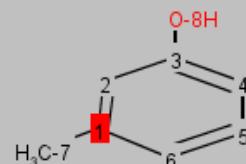
You could bookmark structures if you were logged in!

Details

User input (2D)

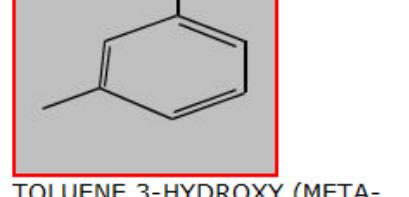
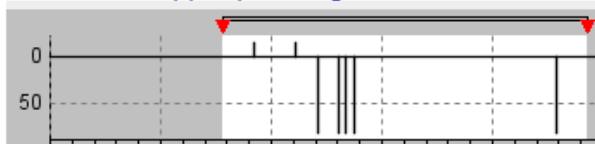
Spectral Data [Additional Data](#)

Type: ^{13}C Double click to get resizable window!



Atom No.	Mult.(coupling const.)	Meas. Shift	expt-0 ocmainz inhouse-2	expt-1
1	S	138.9	139.8	139.3
2	D	116.1	116.2	116.1
3	S	157.7	155.0	154.9
4	D	112.6	112.5	112.7
5	D	129.3	129.4	130.3
6	D	119.8	121.8	122.2
7	Q	21.0	21.1	

Double click upper part to get resizable window!



TOLUENE,3-HYDROXY (META-...)

Example 3: Search by name

- Search by name, literature, CAS Number or formula

Search by Molecule/Spectrum Properties

Search expression: caffeine

Exact

Chemical Name (with Pubchem name resolution)

- Exact
- Regular expression
- Fragment
- Fuzzy

Search by Molecule/Spectrum Properties

Search expression: caffeine

Exact

Chemical Name (with Pubchem name resolution)

Chemical Name (with Pubchem name resolution)

Search

Literature/Author

Cas Number

Chemical Formula

Details

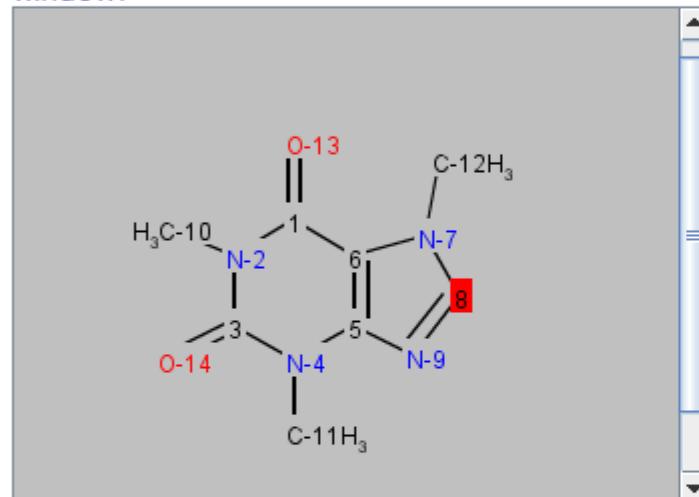
User input (2D) ▾

Show with this coordinate set

Spectral Data Additional Data

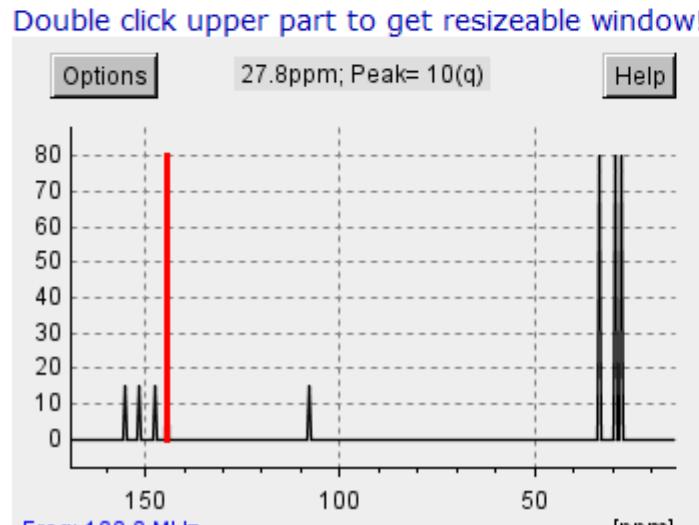
Search for complete spectrum: Similarity measure for the complete spectrum in this record is 100.0.

Type: 13C ⓘ Double click to get resizable window!

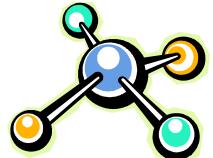


Atom No.	Mult.(coupling const.)	Meas. Shift	Input Shift	Diff. M-I	expt-0 sdbs
1	S	155.3	155.3	0.00	155.32
3	S	151.6	151.6	0.00	148.67
5	S	147.5	147.5	0.00	151.66
6	S	107.8	107.8	0.00	107.51
8	D	144.3	144.3	0.00	141.57
10	Q	27.8	27.8	0.00	29.7
11	Q	29.6	29.6	0.00	27.88
12	Q	33.5	33.5	0.00	33.57

Threshold is 12.50



Search in “Expert Mode”



- Switch to “expert” search mode for additional search options

The diagram illustrates the transition from 'Search by Spectrum' mode to 'Search by Molecule/Spectrum Properties' mode. On the left, a blue rounded rectangle contains the text 'Molecular weight range'. A large blue arrow points from this box towards the right. On the right, there is a screenshot of a search interface. At the top, there is a navigation bar with tabs: Home, Search, Results, Predict, Submit, and a Help icon. Below the navigation bar, the text 'Search by Spectrum' is displayed. Further down, a button labeled 'Switch to expert search mode' is shown. The main search area features a dropdown menu titled 'Search expression:' with the option 'Exact' selected. A list of search properties is provided, with 'Chemical Name (with Pubchem name resolution)' currently highlighted. Other options in the list include Literature/Author, Cas Number, Chemical Formula, Chemical Name (no Pubchem resolution), Chemical Formula (with other elements allowed), Literature>Title, Comment, Canonical Name, Molecule Hyperlink Description, Spectrum Hyperlink Description, Molecule Keyword, Spectrum Category, Multiplicities, Potential C13-Multiplicities, Spectrum NMRShiftDB-Number, Molecule NMRShiftDB-Number, HOSE code, double bond equivalents/smallest set of smallest rings, and Molecular weight (format: from-to). At the bottom right of the interface, the text 'keywords:' is followed by a list: 'α-exomethylene-γ-lactone', 'dkfz spektren database', and 'α-exomethylene-γ-lactones'.



Search by Keyword/Category

- Can search by defined categories
- Can select multiple keywords by holding ctrl key

Search by Keyword/Category

Spectrum categories:

- ab initio
- dkfz spektron database
- HMDB
- NCI
- nmr sharc
- ocmainz inhouse database

Molecule keywords:

- α -exomethylene- γ -lacton
- α -exomethylene- γ -lactones
- α -pyrones
- β -Asarone
- β -hydroxy acetamides
- γ -cembranolide-type
- γ -lactone

Keyword fragment search
 Total keyword search

Search 

Search by Condition

- Search by certain measurement condition or calculated condition

Search by Condition

Measurement conditions:

Temperature [K] 298

Solvent Chloroform-D1 (CDCl₃)

Field Strength [MHz] 125

Assignment Method 1D shift positions

Calculation conditions:

Program Spartan

NMRLocalis HF-GIAO NMR

GeomMethod Hartree-Fock

GeomBasisSet 6-31G*

NMRModel Hartree-Fock

NMRBasisSet 6-31G*

NMRStandard B2H6

Search by condition



Search by selection of isotopes

- Can restrict your search to only certain isotopes

Spectrum type selection

13C

1H

15N

31P

19F

11B

29Si

17O

73Ge

195Pt

33S

Use only reviewed spectra for prediction

Restrict to choice





Example 4: Combining searches

- Search history allows you to repeat and combine previous searches
- For example, combining a substructure with a spectrum search

Search History				
Repeat	Type of search	Mode	Value	Hits
<input type="checkbox"/>	subspectr um search	--	11.345651 128.42825 ...	300
<input type="checkbox"/>	chemical name using Pubchem	fragment	InChI=1/C21H22N2O2 fragment/c24-18-10-16-19-13-9-17-21(6-7-22(17)11-12(13)5-8-25-16)14-3-1-2-4-15(14)23(18)20(19)21 2/h1-5,13,16-17,19-20H,6-11H	300
<input checked="" type="checkbox"/>	substructure search	--		300
<input checked="" type="checkbox"/>	total similarity spectrum search	--	11.345651 26.956678 ...	300

Combined search mode

not and or
Repeat marked search(es)





Example 5: Prediction

- Structure search can find a spectrum only for those compounds in the database
- For any structure, you can *predict* the spectrum

Predict

Predict an NMR Spectrum

Help

Edit View Insert Report

Chemical structure input area showing a 2D chemical structure of 4-methyl-2-(4-methylphenyl)-2H-pyridine-4-ol.

Table of predicted NMR shifts:

No.	Shift	deviation	spheres	min	max	mean	median	values
1	126.36	2.98	3	117.40	134.30	126.36	126.85	172
2	128.47	3.30	4	123.30	135.90	128.47	127.05	134
3	135.26	6.25	2	121.80	154.80	135.26	135.90	302
4	125.85	0.21	5	125.70	126.00	125.85	125.85	2
5	141.30	0.42	3	141.00	141.60	141.30	141.30	2
6	39.17	♦	2	39.17	39.17	39.17	39.17	1
7	128.15	0.21	4	128.00	128.30	128.15	128.15	2
8	25.40	0.14	4	25.30	25.50	25.40	25.40	2
9	44.87	♦	3	44.87	44.87	44.87	44.87	1
10	15.30	0.00	5	15.30	15.30	15.30	15.30	2

Get mol file

Structure file: Browse... Import file

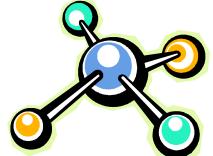
Import text by copy & paste Import from structures history

Use measured and/or calculated spectra

¹³C Predict spectrum

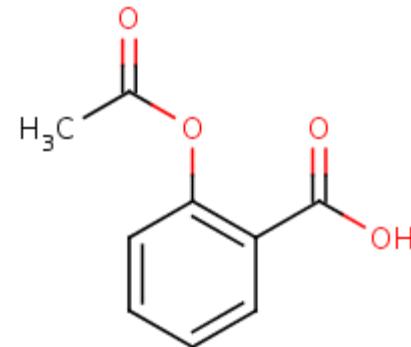


Time for exercises



Exercise 1

- Given the following chemical structure can you **identify** whether its spectrum is in the database? (Hint use the chemical structure search, use ChEBI to find and import the structure)



- Using the same chemical structure can you **predict** the spectrum?

Exercise 2

1. Search for the term ‘aspirin’ using NMRShiftDB? Do you find any spectrum?

3. Using the answer in 2.1 can you name any synonyms of aspirin?

Exercise 2

- Given the following spectrum data, can you find **subspectrums** which match it?

122

132

126

135

124

151

170



To summarise

- NMRShiftDB supports structure identification.
- NMRShiftDB can be used to identify chemicals by searching using spectrum data which you have measured
- NMRShiftDB can be used to search by chemical structure, names and additional properties
- NMRShiftDB can be used to predict spectrum for a specific chemical structure if its not found in the database.

Conclusions

- You can use the data for your research.
- Since NMRShiftDB is a community effort, you can contribute yourself using the “Submit” function.
- The Specclipse application available at <http://sourceforge.net/projects/nmrshiftdb/> offers a user-friendly client for the database.

Thank you

