



NMR-based identification of specialized metabolites

Jean-Marc Nuzillard

*Institut de Chimie Moléculaire de Reims (ICMR), UMR CNRS 7312, SFR CAP Santé,
Université de Reims Champagne-Ardenne, France*

jm.nuzillard@univ-reims.fr

Analysis of organic compound mixtures

- Steadily increasing interest for sources of renewable raw materials from **plants**
- Medicinal drugs from **plants**
- Cosmetic ingredients from **plants**
- **But:** Plants rarely produce organic compounds in chemically pure state.
- **Mixture analysis plays a central role in plant chemistry.**

Quick identification of
known compounds

=

Structural dereplication

=

More time devoted to the
study of **new** compounds



« CAMEL » WORKFLOW

- Extract fractionation by Centrifugal Partition Chromatography (CPC)
→ 10 to 15 fractions
- ^{13}C RMN spectra of fractions
- “Binning” of NMR spectra (1200 bins, each 0.2 ppm wide)
- Intensity Table: each row is related to the chemical shift value at the center of a bin and each column is related to a chromatographic fraction. Empty rows are removed from the Table.
- Row permutations in order to group chemical shift values for which chromatographic profiles are similar
- Determination of chemical shift clusters that can be assigned to a single compound
- **Compound identification by database querying**

- **Centrifugal Partition Chromatography**
- Partition of analytes between two liquid phases
- The column contains a high number of connected partition cells
- The stationary phase is maintained inside the column by action of the centrifugal force
- Analytes are injected in the first partition cells
- The mobile phase percolates through the stationary phase
- **No irreversible adsorption on a solid stationary phase**
- Isocratic or graduated elution, “pH-zone refining”, ion exchange
- All what goes in finally comes out, from one side or the other
- **High flow rates, 20 mL/min**
- Injection of 5g in a 200 mL inner volume column

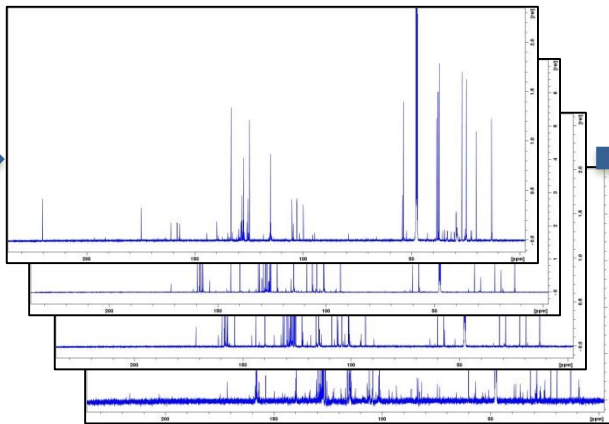
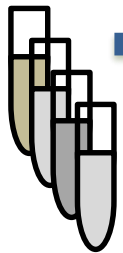
- **CPC is a preparative method**



- Fractions from CPC are analyzed by ¹³C NMR
- One carbon atom, one peak (Unless symmetry or accident occurs)
- Low probability of peak superimposition
- Low sensitivity, presumably
- 600 MHz, cryoprobe, cooled ¹³C coil
- ¹H NMR
 - Not always enough information there
 - Complex spectra due to homonuclear couplings
- Other choices:
 - Pure shift ¹H 1D NMR (difficult...)
 - HSQC
 - HMBC

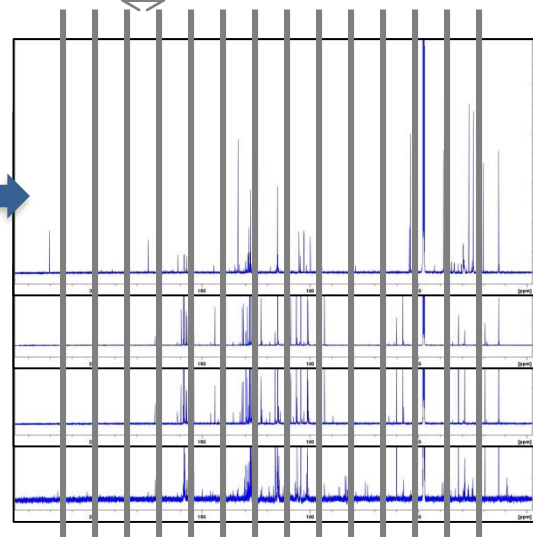
CREATION OF THE NMR DATA TABLE

Fractions



Analysis by ^{13}C NMR

0.2 ppm

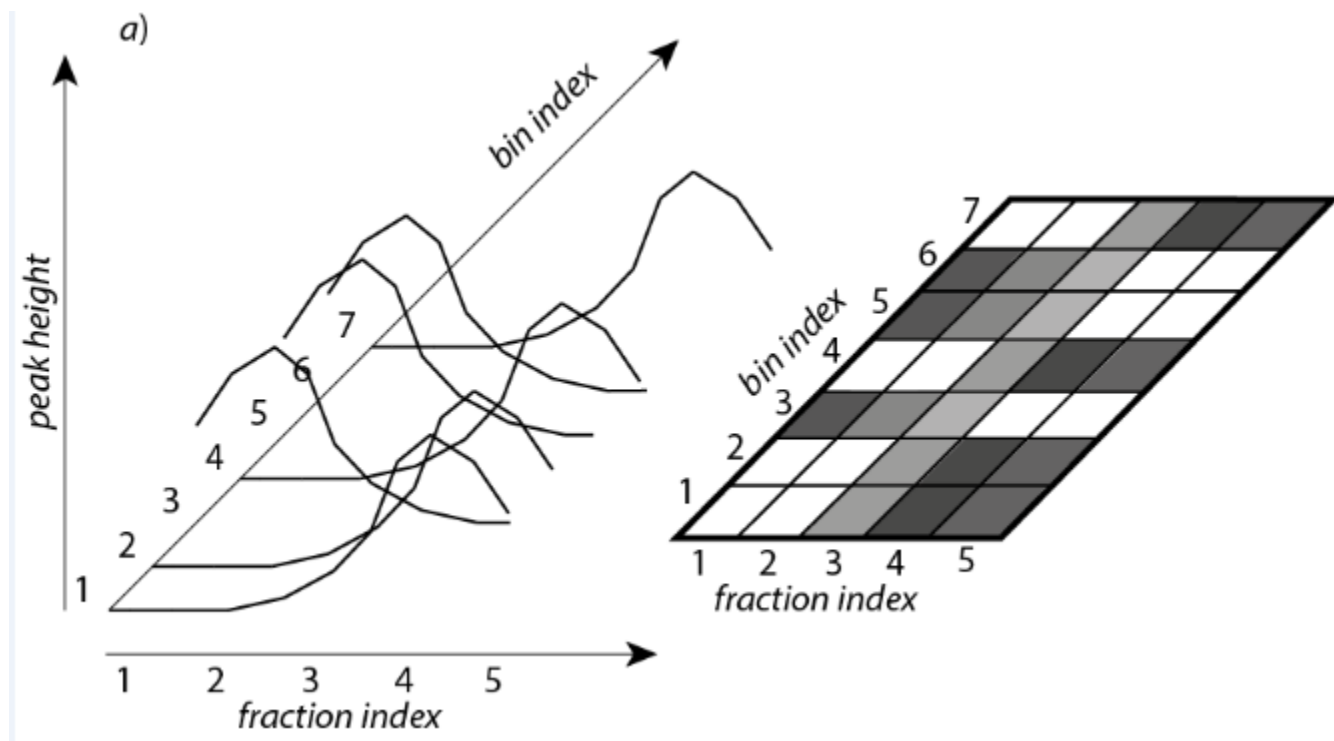


ppm	f1	f2	...	fx
16.3	2E+08	1E+08	...	0
17.5	0	0	...	0
18.7	1E+08	1E+08	...	0
⋮	⋮	⋮	...	⋮
Signal intensities				
176.1	3E+07	0	...	0
177.7	0	0	...	4E+07
177.9	6E+07	6E+07	...	0
199.5	7E+07	5E+07	...	0

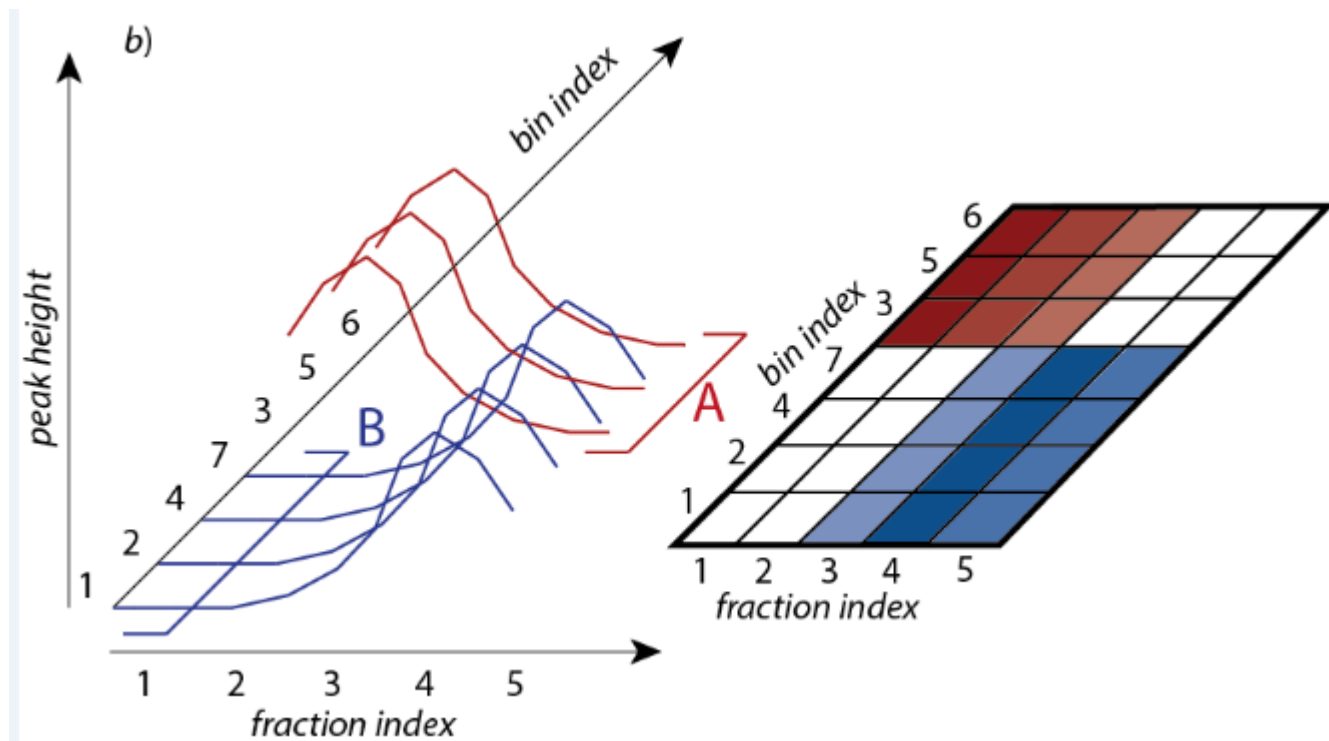
- Automatic peak picking
- File conversion
- Alignment



SIMPLIFIED RAW DATA TABLE

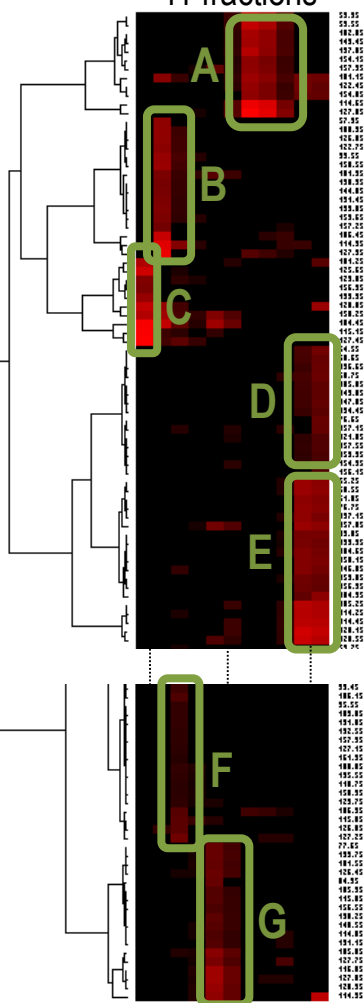


REORDERED LINES IN DATA TABLE BY MEANS OF HIERARCHICAL CLUSTER ANALYSIS (HCA)



REORDERED LINES IN DATA TABLE BY MEANS OF HIERARCHICAL CLUSTER ANALYSIS (HCA)

11 fractions

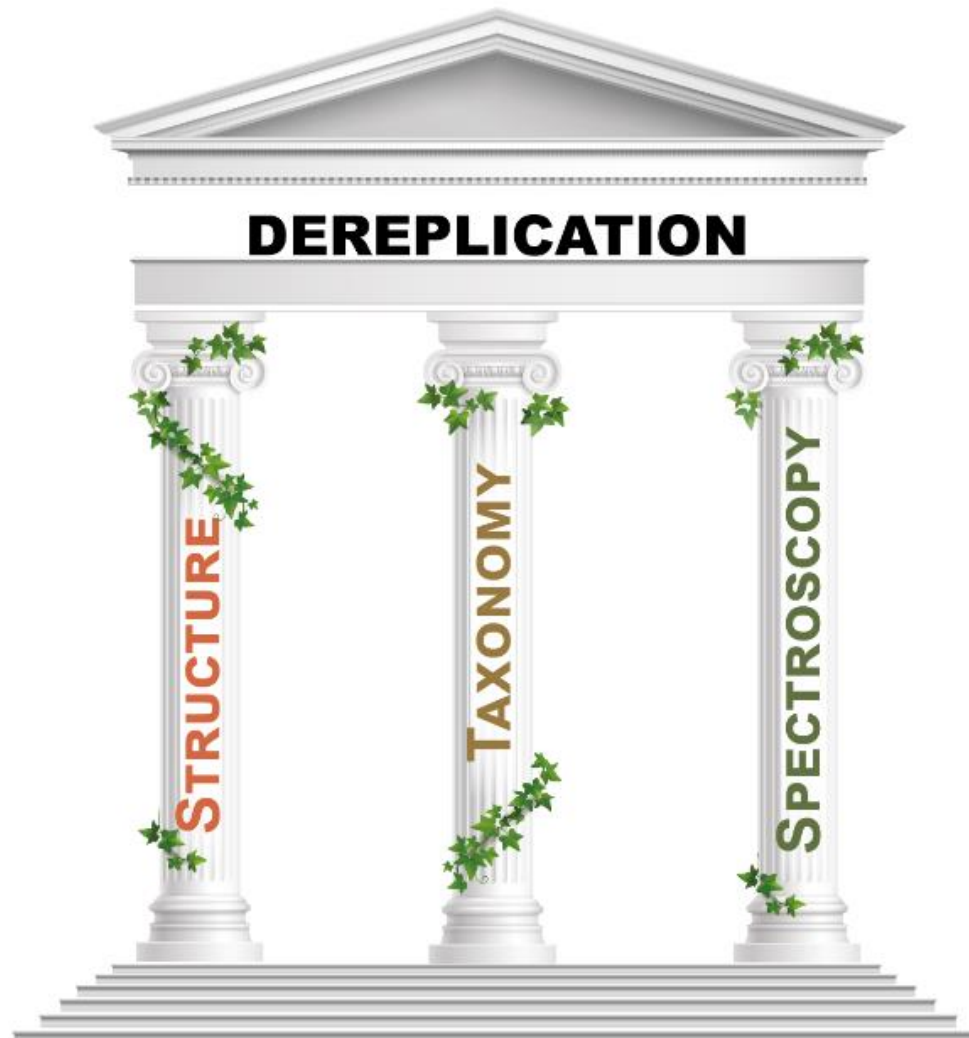


} Group of chemical shift values that can be assigned to a **pure compound**

- If a part of the list of chemical shifts is the one related to a **known compound**, how is it possible to find its structure?
- **The same question holds for a compound that was isolated in pure state.**
- Answering this question requires the availability of *ad hoc* **databases**.



**THE THREE PILLARS OF
STRUCTURAL DEREPLICATION**





STRUCTURE (+ TAXONOMY)

- Considering only the set of known **natural** products improves search efficiency (300k?). **Where to find them?**
- **ISDB** zenodo.org/record/5607264 About 200k structures + MS
- **KNApSAcK** www.knapsackfamily.com/knapsack_core/top.php was a first possibility. About 50k compounds. Closed source.
- **COCONUT** coconut.naturalproducts.net/ has more compounds than KS but not all are natural (407k unique compounds). Open source.
- **LOTUS** lotus.naturalproducts.net/ is like COCONUT but more selective in its compound selection (276k compounds). It includes natively biological and chemical taxonomies, as well as links to Wikidata www.wikidata.org/wiki/. Open source.



STRUCTURE + ^{13}C NMR

- The **experimental** ^{13}C NMR data are **scattered** in chemical literature, are **not exhaustive**, and sometimes of **low reliability**.
- **Predicted** data are a good alternative to experimental data.
- Prediction requires quick, autonomous, and reliable predictors.
- 200.000 seconds = 2 days, 7 hours, 33 minutes and 20 seconds
- **Solution 1** : nmrshiftdb.nmr.uni-koeln.de with the Java code that allows for predictions locally. Very fast.
- **Solution 2** : www.acdlabs.com ACD/Labs software and
 - “undocumented use” of the experimental chemical shift validation tool (quick, unattended, 2 compounds per second)
 - accurate predictor (slow, GUI-driven, 1 minute per compound).



TWO DEREPLICATION STRATEGIES

➤ **Initially Untargeted:**

- The biggest possible collection of structures is associated with predicted spectroscopic data and with taxonomic information. This has to be carried out only at the time of database creation.
- The results of a structure search can be refined according to taxonomic criteria, if needed.

➤ **Initially Targeted:**

- The structures are first selected according to taxonomic criteria. The **spectroscopic data** are then predicted **on demand** on a small sized set of structures.
- Querying a targeted database is quick and provides pertinent structure proposals.



INITIALLY TARGETED STRATEGY

- **KnapsackSearch (KS)** github.com/nuzillard/KnapsackSearch
 - Provide a list of genera related to an organism family to KS
 - Wikipedia is your friend
 - A Python script
 - Queries KNAPSAcK through the web about each genus
 - Collects all found structures related to a family
 - Predicts chemical shifts with nmrshiftdb2
 - Creates a .sdf file ACD/Labs software can read
 - Requires an internet connection (slow)
 - Relies on the writing style of data sent back by the KNAPSAcK web site (coding in HTML). Will not last forever.
 - Relies on RDKit – <https://www.rdkit.org/>



INITIALLY TARGETED STRATEGY

- **VersaDB** https://github.com/simremy/versadb_tk
 - Developed by Simon Remy and Julien Cordonnier, @URCA
 - GUI software
 - Compound selection by biological and chemical taxonomy
 - Queries LOTUS through the Internet
 - RMN data prediction by nmrshiftdb2
 - MS² data prediction by CFM-ID 4.0
 - Relies on a connection to the Internet
 - Relies on RDKit
 - Should be published soon...



INITIALLY UNTARGETED STRATEGY

- History : SISTEMAT Knowledge Base, Pr. VdP Emerenciano, Brazil
 - Beginning of the 1990's.
 - The Three Pillars of Dereplication were already there
 - Closed source databases and software

- « Predicted NMR of Natural Products » (PNMRNP), @ICMR
 - Starting from ISDB and NMR-supplemented using nmrshiftdb2
 - Proof of principle, practical use is not recommended.



INITIALLY UNTARGETED STRATEGY

➤ **ACD_LOTUS**

- Structures from LOTUS, supplemented with predicted NMR data
 - Uses ACD/Labs quick validation algorithm for predictions
- Uses RDKit and the tools developed for KS et PNMRNP
- 218.478 structures in LOTUS v9
- zenodo.org/record/7124055 .sdf zipped file
- Structures and spectra merged into nmrshiftdb2 by Stefan Kuhn
- The publicly available .sdf file can be imported in ACD/Labs
 - Using **acd_lotusv9.NMRUDB** does not required a connection to the Internet



EXAMPLE, ACD_LOTUS

- Structure selection of a known specialized metabolite
- Copy the list of the associated ^{13}C NMR chemical shift values
- Look for matching compounds by means of the
 - ACD/Labs database search tool
 - nmrshiftdb2 search tool

EXAMPLE, ACD_LOTUS – ELLIPTICINE

- Ellipticine in the Biological Magnetic Resonance data Bank (BMRB)
- bmr.io/metabolomics/metabolomics_standards.php?dataset=metabolomics

Filter by dataset

NMRFAM highly curated metabolites

Go

FAM : Facility at Madison (Wisconsin)

Jump to molecules beginning with: A B C D **E** F G H I J K L M N O P Q R S T U V X Z

6-ethylmercaptapurine	(E)-3,7,11,15-tetramethylhexadec-2-en-1-ol	(E)-6-(4-hydroxy-6-methoxy-7-methyl-3-oxo-1H-2-benzofuran-5-yl)-4-methylhex-4-enoic acid
(~)-Eburnamonine	EGTA	elaidic acid
Ellagic acid	Ellipticine	Emodin
(-)-Epicatechin	epichlorohydrin	Epinephrine
epsilon-caprolactam	Epsilon-caprolactone	ergocalciferol

Ellipticine bmse001340 - Data

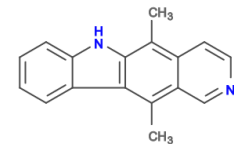
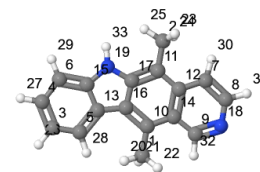
Ellipticine synonyms

Wikipedia: Ellipticine

Go

Search Wikipedia

Graphical representations:





FOR WOOD AND NMR LOVERS

➤ https://bmrbl.io/metabolomics/metabolomics_standards.php?dataset=NMR+Database+of+Lignin+and+Cell+Wall+Model+Compounds

Filter by dataset

A description of these data, collected by Sally Ralph and Larry Landucci at the USDA Forest Products Lab, and John Ralph's group at the USDA Dairy Forage Research Center and (more recently) the UW Biochemistry Department, is available as a [pdf file](#). This includes a description of the naming conventions used as well as experimental conditions and a structure index.

Jump to molecules beginning with: **A B C D E F G H I L M O P S T V X**

A		
1-(4-acetoxy-3,5-dimethoxyphenyl)-1,3-diacetoxy-2-(2-methoxyphenoxy)propane	1-(4-acetoxy-3,5-dimethoxyphenyl)-1,3-diacetoxy-2-[4-(1-acetoxyethyl)-2,6-dimethoxyphenoxy]propane	1-(4-acetoxy-3,5-dimethoxyphenyl)-1,3-diacetoxy-2-[4-(1-acetoxyethyl)-2-methoxyphenoxy]
1-(4-acetoxy-3,5-dimethoxyphenyl)-1,3-diacetoxy-2-[4-(1-acetoxyethyl)phenoxy] propane	1-(4-acetoxy-3,5-dimethoxyphenyl)-2-(4-acetoxy-2-methoxyphenoxy)-3-hydroxypropan-1-one	1-(4-acetoxy-3,5-dimethoxyphenyl)-2-(4-acetyl-2,6-dimethoxyphenoxy)-3-hydroxypropan-1-one
1-(4-acetoxy-3,5-dimethoxyphenyl)-2-(4-acetylphenoxy)-3-hydroxypropan-1-one	1-(4-acetoxy-3,5-dimethoxyphenyl)-3-hydroxy-2-(2-methoxyphenoxy)propan-1-one	2-(4-Acetyl-2,6-dimethoxyphenoxy)-1-(3,4,5-trimethoxyphenyl)ethanone



Set 1

Sample: 5mg in DMSO, ref: TMS
Conditions: temperature: 298K
Spectrometer: Bruker Avance III - 500MHz

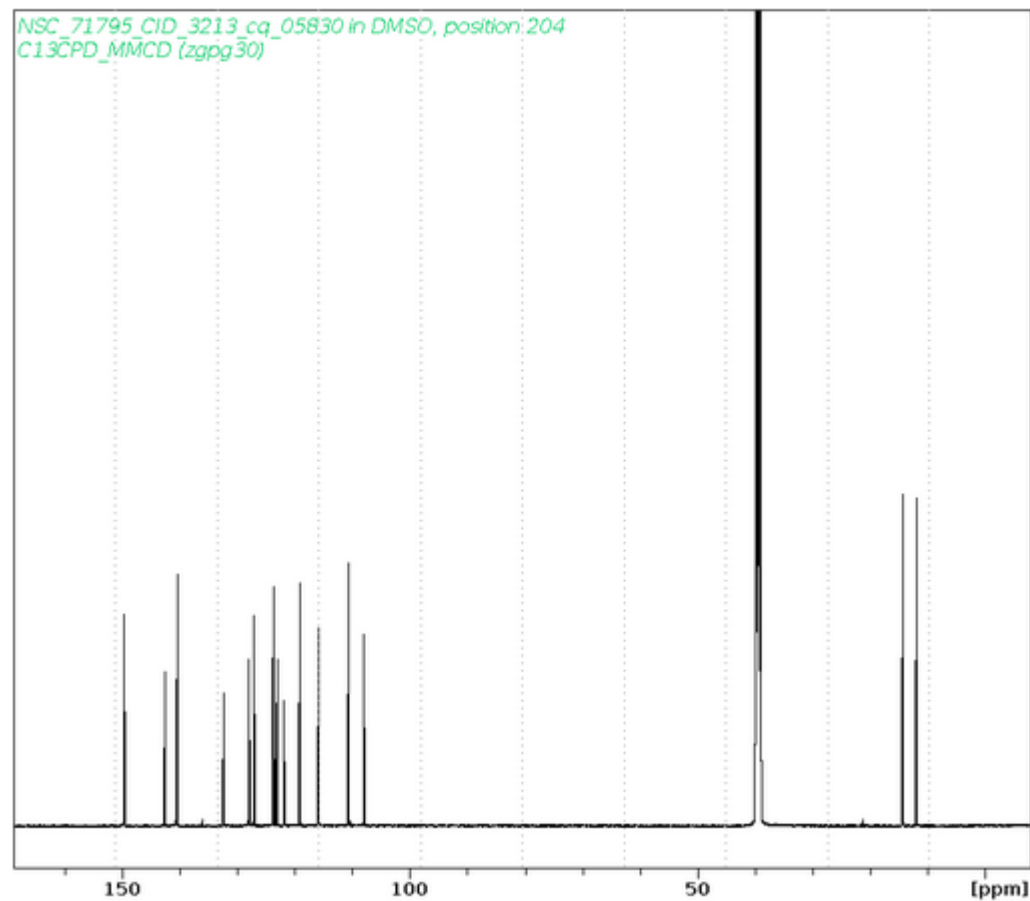
Atom ID	Value	Ambiguity Code
C1	14.3692	1
C2	11.9824	1
C3	119.191	4
C4	127.1373	4
C5	123.8352	4
C6	110.7067	4
C7	115.8997	1
C8	140.5188	1
C9	149.7257	1
C10	128.066	1
C11	132.4729	1
C12	108.0425	1
C13	123.1352	1
C14	121.9748	1
C15	142.6759	1
C16	123.3902	1
C17	140.5486	1

Experimental values

BACK TO ELLIPTICINE

3: 1D 13C

Sample: 5mg in DMSO, ref: TMS
Conditions: temperature: 298K
Spectrometer: Bruker Avance III - 500MHz



Download time domain data: [bmse001340_3.zip](#)

ELLIPTICINE

Query

14.3692
11.9824
119.1910
127.1373
123.8352
110.7067
115.8997
140.5188
149.7257
128.0660
132.4729
108.0425
123.1352
121.9748
142.6759
123.3902
140.5486

Open file `acd_lotus9.NMRUDB`

ACD/C-H NMR Predictors and DB: Database Window - [C:\USERS\JMN\DOCUMENTS\CNRS22\LOTUSV9_ACD\ACD_LOTUSV9.NMRUDB]

Database View Record Search Lists Table Training Options ACD/Labs Help

Atom No. 13C Shift 1H Shift

Atom No.	13C Shift	1H Shift
1	23.02	
2	28.17	
3	22.76	
4	39.74	
5	24.06	
6	36.46	
7	35.97	
8	18.86	
9	56.56	
10	28.36	
11	24.48	
12	56.91	
13	32.16	
14	32.12	
15	121.49	
16	141.08	
17	42.44	
18	71.59	
20	31.72	
21	37.55	
22	36.67	
23	19.46	
24	50.52	
25	21.29	
26	40.07	

Formula: $C_{27}H_{46}O$
FW: 386.6535
SMILES_ID: Q43656_2
Predicted 13C shifts: 1[1] 23.02 ; 2[2] 28.17 ; 3[3] 22.76 ; 4[4] 39.74 ; 5[5] 24.06 ; 6[6] 36.46 ; 7[7] 35.97 ; 8[8] 18.86 ; 9[9] 56.56 ; 10[10] 28.36 ; 11[11] 24.48 ; 12[12] 56.91 ; 13[13] 32.16 ; 14[14] 32.12 ; 15[15] 121.49 ; 16[16] 141.08 ; 17[17] 42.44 ; 18[18] 71.59 ; 20[20] 31.72 ; 21[21] 37.55 ; 22[22] 36.67 ; 23[23] 19.46 ; 24[24] 50.52 ; 25[25] 21.29 ; 26[26] 40.07

ID: 1 A: 1/218478 B: 218478 Last Updated: 28/09/2022 16:27 Exclude from Calculations Single DB

1-ChemSketch 2-CNMR Spectrum 3-HNMR Spectrum 4-History 5-Database

ELLIPTICINE

Search by Chemical Shifts

Enter CNMR Query Shifts (example: 32.0 128.1..130)

14.3692
11.9824
119.1910
127.1373
119.2357

CNMR Shifts Looseness Factor (+/-): 4.00

Minimum Number of CNMR Query Shifts to Match: 17

Enter HNMR Query Shifts (example: 3.2 6.1..6.4 10s[10])

HNMR Shifts Looseness Factor (+/-): 0.00

Minimum Number of HNMR Query Shifts to Match:

Search through Unambiguous Assignment
 Do not Match One Chemical Shift for Several Shift Queries

Do not Sort Result
 Sort Result by HQI Based on Minimal Distances
 Sort Result by HQI Based on Shifts Distribution

OK Cancel Help

Search Message

1690 hits found for your query

OK

ACD/C+H NMR Predictors and DB: Database Window - [C:\USERS\JMN\DOCUMENTS\CNRS22\LOTUSV9_ACD\ACD_LOTUSV9.NMRUDB]

Database View Record Search Lists Table Training Options ACD/Labs Help

Atom No.	13C Shift	1H Shift
1	12.0	
2	111.99	
3	132.45	
4	115.9	
5	140.5	
7	149.7	
8	122.0	
9	128.1	
10	14.4	
11	125.75	
12	140.55	
14	142.7	
15	110.7	
16	127.15	
17	119.2	
18	123.8	
19	122.55	

Formula: C₁₇H₁₄N₂
FW: 246.3065
SMILES_ID: Q10359556_90835
Predicted 13C shifts: 1[1] 12.00 ; 2[2] 111.99 ; 3[3] 132.45 ; 4[4] 115.90 ; 5[5] 140.50 ; 7[7] 149.70 ; 8[8] 1

Wikidata Q-Id: Q10359556

ELLIPTICINE IN WIKIDATA

The image shows a screenshot of the Wikidata Main Page in a Firefox browser. The browser's address bar shows the URL https://www.wikidata.org/wiki/Wikidata:Main_Page. The page features a navigation menu at the top with options like 'Main Page', 'Discussion', 'Read', 'View source', and 'View history'. A search bar in the top right corner contains the text 'Q10359556', which is circled in red. A dropdown menu below the search bar displays the search results: 'ellipticine (Q10359556) chemical compound' and 'Search for pages containing Q10359556'. The main content area is a 'Welcome to Wikidata' message, stating it is 'the free knowledge base with 102,507,809 data items that anyone can edit.' Below this message are two columns of information: 'Welcome!' and 'Learn about data'. The 'Welcome!' section describes Wikidata as a free and open knowledge base. The 'Learn about data' section encourages users to develop and improve their data literacy. The background of the page features a network diagram with nodes and connecting lines, and a decorative image of petri dishes with green cultures.

Fichier Édition Affichage Historique Marque-pages Outils Aide

Wikidata

← → ↻ 🏠 🔒 🔑 🔍 ☆ 📧 📧 📧 📧 📧 📧

https://www.wikidata.org/wiki/Wikidata:Main_Page

Débuter avec Firefox

Autres marque-pages

English Not logged in Talk Contributions Create account Log in

Main Page Discussion Read View source View history

ellipticine (Q10359556)
chemical compound
Search for pages containing Q10359556

collaborative

open

free

multilingual

structured

Welcome to Wikidata

the free knowledge base with 102,507,809 data items that anyone can edit.

[Introduction](#) • [Project Chat](#) • [Community Portal](#) • [Help](#)

Want to help translate? [Translate the missing messages.](#)

Welcome!

Wikidata is a free and open knowledge base that can be read and edited by both humans and machines.

Wikidata acts as central storage for the **structured data** of its Wikimedia sister

Learn about data

New to the wonderful world of data? [Develop and improve your data literacy through content](#) designed to get you up to speed and feeling comfortable with the fundamentals in no time.

Main page
Community portal
Project chat
Create a new Item
Recent changes
Random Item
Query Service
Nearby
Help
Donate

Lexicographical data
Create a new Lexeme
Recent changes
Random Lexeme

Tools
[What links here](#)
[Related changes](#)
[Special pages](#)
[Permanent link](#)

ELLIPTICINE

Eichier Édition Affichage Historique Marque-pages Outils Aide

ellipticine - Wikidata

https://www.wikidata.org/wiki/Q10359556

Autres marque-pages

English Not logged in Talk Contributions Create account Log in

Item **Discussion** Read View history Search Wikidata

ellipticine (Q10359556)

chemical compound [edit](#)

NSC-71795 | Elliptisine | NSC 71795 | ellipticines

[In more languages](#)

Statements

instance of	chemical compound edit
	0 references + add reference
	carbazole alkaloid edit
	1 reference
	type of chemical entity edit
	0 references

ELLIPTICINE

found in taxon



Ochrosia elliptica

edit

▼ 5 references

stated in

The biogenetic, synthetic and biochemical aspects of ellipticine, an antitumor alkaloid

stated in

Fluorodensitometric assay of 6H-pyrido[4,3-b]-5,11-dimethylcarbazoles (ellipticine and derivatives) biosynthesized by Ochrosia elliptica cultures in vitro

stated in

Antitumor alkaloids in callus cultures of Ochrosia elliptica

stated in

Alkaloid Production by Ochrosia elliptica Cell Suspension Cultures

stated in

Alkaloids of Ochrosia elliptica Labill.1

+ add reference



Ochrosia moorei

edit

▶ 1 reference

ELLIPTICINE



Item [Discussion](#)

Alkaloids of *Ochrosia elliptica* Labill.1 (Q104937747)

scientific article published on 10 March 2005

Identifiers

DOI

10.1021/JA01517A031 CR

edit

▼ 0 references

+ add reference


+ add value


+ add statement


ELLIPTICINE

Alkaloids of *Ochrosia elliptica* Labill.¹

Sidney Goodwin, A. F. Smith, and E. C. Horning

Cite this:  *J. Am. Chem. Soc.* 1959, 81, 8, 1903–1908

Publication Date: April 1, 1959 

<https://doi.org/10.1021/ja01517a031> 

© American Chemical Society

[RIGHTS & PERMISSIONS](#)  Subscribed

 ACS Legacy Archives

Article Views

513

Altmetric


3

Citations

202

[LEARN ABOUT THESE METRICS](#)



PDF (788 KB) 

open URL 

April 20, 1959

ALKALOIDS OF *Ochrosia elliptica* LABILL.

1903

[CONTRIBUTION FROM THE LABORATORY OF CHEMISTRY OF NATURAL PRODUCTS, NATIONAL HEART INSTITUTE, NATIONAL INSTITUTES OF HEALTH]

Alkaloids of *Ochrosia elliptica* Labill.¹

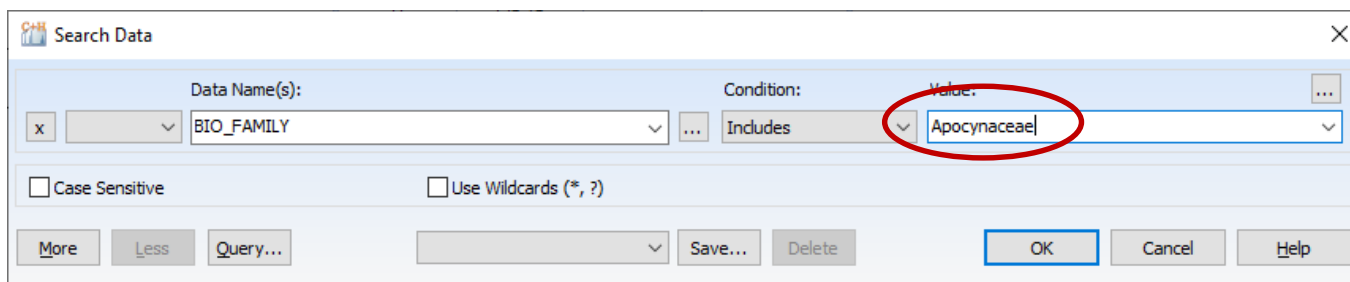
BY SIDNEY GOODWIN, A. F. SMITH² AND E. C. HORNING

RECEIVED JULY 28, 1958

Four alkaloids have been isolated from leaves of *Ochrosia elliptica* Labill. One has been shown to be identical with isoreserpiline. The other three, ellipticine, methoxyellipticine and elliptinine, have been characterized and certain features of their structures have been suggested.

ELLIPTICINE – INITIALLY TARGETED STRATEGY

Apocynaceae alkaloids

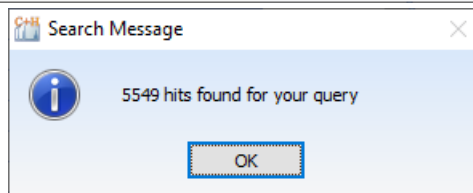


Search Data

Data Name(s):	Condition:	Value:
<input checked="" type="checkbox"/> BIO_FAMILY	Includes	Apocynaceae

Case Sensitive Use Wildcards (*, ?)

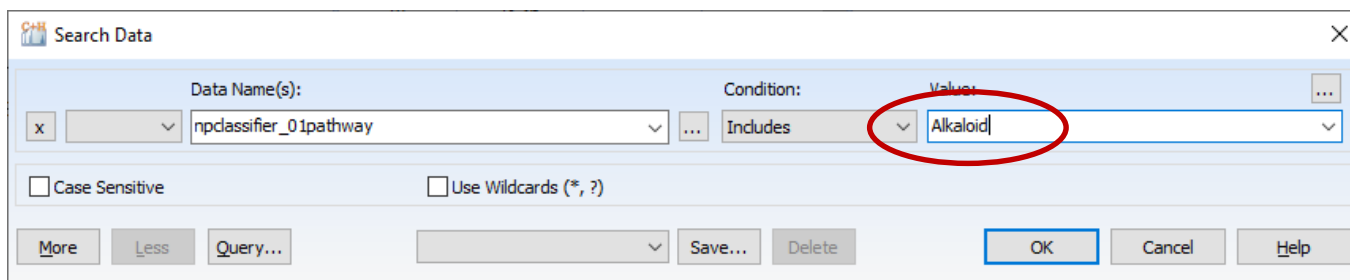
More Less Query... Save... Delete OK Cancel Help



Search Message

5549 hits found for your query

OK

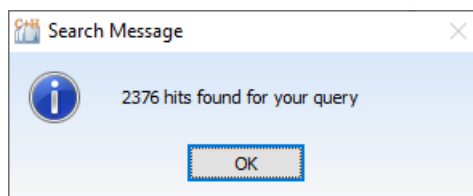


Search Data

Data Name(s):	Condition:	Value:
<input checked="" type="checkbox"/> npclassifier_01pathway	Includes	Alkaloid

Case Sensitive Use Wildcards (*, ?)

More Less Query... Save... Delete OK Cancel Help



Search Message

2376 hits found for your query

OK

ELLIPTICINE

Search by Chemical Shifts

Enter CNMR Query Shifts (example: 32.0 128.1..130)

14.3692
11.9824
119.1910
127.1373
122.555

CNMR Shifts Looseness Factor (+/-): 4.00

Minimum Number of CNMR Query Shifts to Match: 17

Enter HNMR Query Shifts (example: 3.2 6.1..6.4 10s[10])

HNMR Shifts Looseness Factor (+/-): 0.00

Minimum Number of HNMR Query Shifts to Match:

Search through Unambiguous Assignment

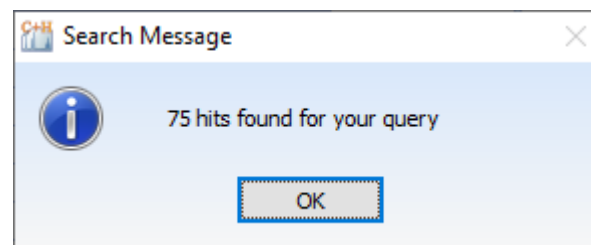
Do not Match One Chemical Shift for Several Shift Queries

Do not Sort Result

Sort Result by HQI Based on Minimal Distances

Sort Result by HQI Based on Shifts Distribution

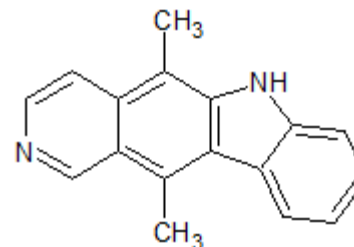
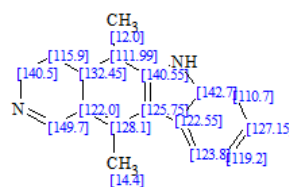
OK Cancel Help



ACD/C+H NMR Predictors and DB: Database Window - [C:\USERS\JMN\DOCUMENTS\CNRS22\LOTUSV9_ACD\ACD_LOTUSV9.NMR\UDB]

Database View Record Search Lists Table Training Options ACD/Labs Help

Atom No.	¹³ C Shift	¹ H Shift
1	12.0	
2	111.99	
3	132.45	
4	115.9	
5	140.5	
7	149.7	
8	122.0	
9	128.1	
10	14.4	
11	125.75	
12	140.55	
14	142.7	
15	110.7	
16	127.15	
17	119.2	
18	123.8	
	122.55	

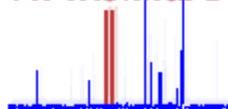


EXAMPLE, NMRSHIFTDB2 – ELLIPTICINE

Query

14.3692
11.9824
119.1910
127.1373
123.8352
110.7067
115.8997
140.5188
149.7257
128.0660
132.4729
108.0425
123.1352
121.9748
142.6759
123.3902
140.5486

NMRShiftDB



Current usage is:

Registered Users: 2138

Structures: 258158

Spectra: Measured 53954, calculated 396566

Impressum

[Home](#) [Search](#) [Results](#) [Quick Check](#) [Predict](#) [Assignment](#)

Search by Spectrum

Switch to expert search mode

Browse all structures

Input list ⓘ: [Input format](#)

14.3692
11.9824
119.1910
127.1373
123.8352
110.7067
115.8997
140.5188
149.7257
128.0660
132.4729

Spectrum type to search:

13C ▼

Subspectrum

Complete

Search by spectrum

ELLIPTICINE

total similarity -- 13C
 spectrum search 14.3692|11.9824|119....

[Spectral Data](#) [Additional Data](#) [Download](#)

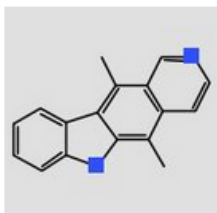
Results: 300

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

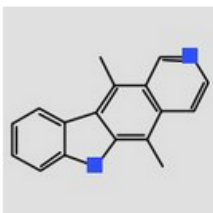
Search for complete spectrum: Similarity measure for the complete spectrum in this record is 92.92.

Type: 13C

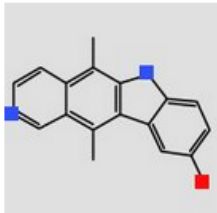
Next structure



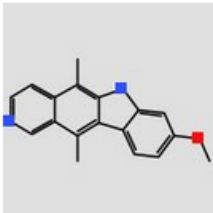
Similarity: 90.15 %



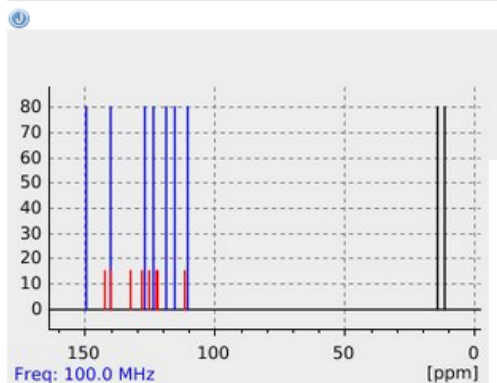
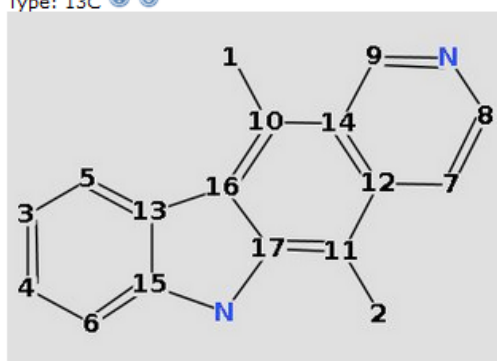
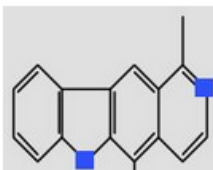
Similarity: 82.55 %



Similarity: 71.02 %

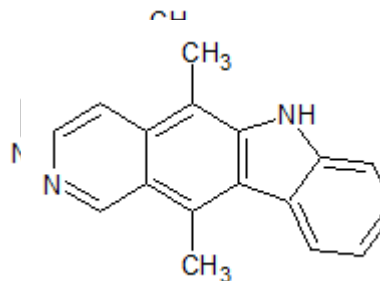


Similarity: 69.05 %



Atom	Mult.(coupling const.)	Prediction Shift	Input Shift	Diff. M-I	Prediction 0 hose_lotusv7
1	Q	14.40	14.3692	0.03	14.32
2	Q	12.00	11.9824	0.02	11.6
3	D	119.20	119.191	0.01	120.19
4	D	127.15	127.1373	0.01	125.25
5	D	123.80	123.3902	0.41	123.06
6	D	110.70	108.0425	2.66	111.2
7	D	115.90	115.8997	0.00	115.7
8	D	140.50	140.5188	0.02	142.2
9	D	149.70	149.7257	0.03	150.2
10	S	128.10	128.066	0.03	128.34
11	S	111.99	110.7067	1.28	111.0
12	S	132.45	132.4729	0.02	130.38
13	S	122.55	123.1352	0.59	121.95
14	S	122.00	121.9748	0.03	125.6
15	S	142.70	142.6759	0.02	140.95
16	S	125.75	123.8352	1.91	124.83
17	S	140.55	140.5486	0.00	136.36

Threshold is 5.88





TAKE HOME MESSAGE

- The SDF-formatted `acd_lotusv9` database is available from <https://zenodo.org/record/7124055>
- The `acd_lotusv9.sdf` file can be imported as a searchable database file in ACD/Labs software products
- The structures and chemical shift values from `acd_lotusv9` are already searchable in <https://nmrshiftdb.nmr.uni-koeln.de/>
- The assessment of the known compound identification process was carried out on a small set of 58 secondary metabolites. The results of this study should be published soon.



Natural Products Chemistry Team @ICMR

