NMR data reporting for better article reviewing?

Jean-Marc Nuzillard [a] **and Christoph Steinbeck** [b]

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

[b] *Cheminformatics and Computational Metabolomics, Friedrich-Schiller-University, Jena, Germany*

THE STARTING POINT

Phytochemistry 142 (2017) 38–50

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Polyprenylated acylphloroglucinols from *Hypericum scabrum*

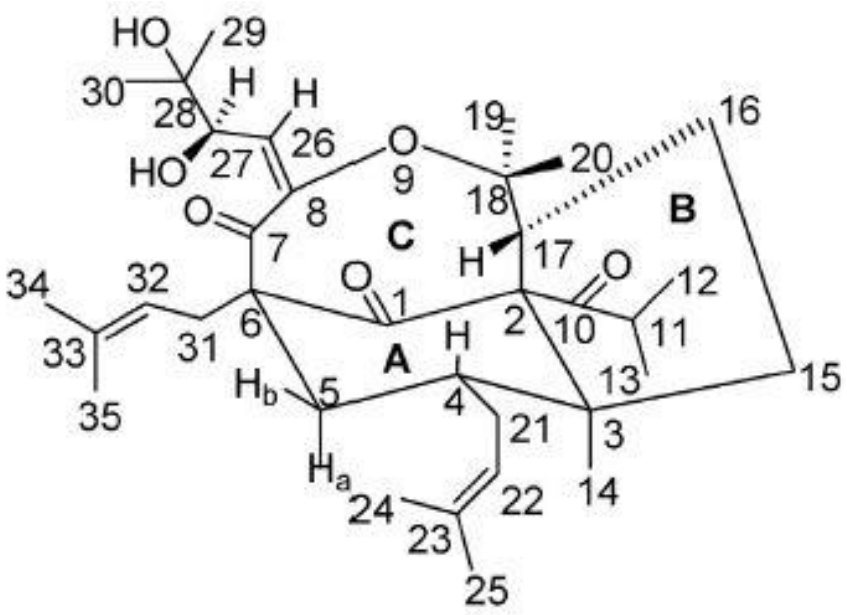
Rangdong Liu, Yalun Su, Jianbo Yang*, Aiguo Wang**

State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China

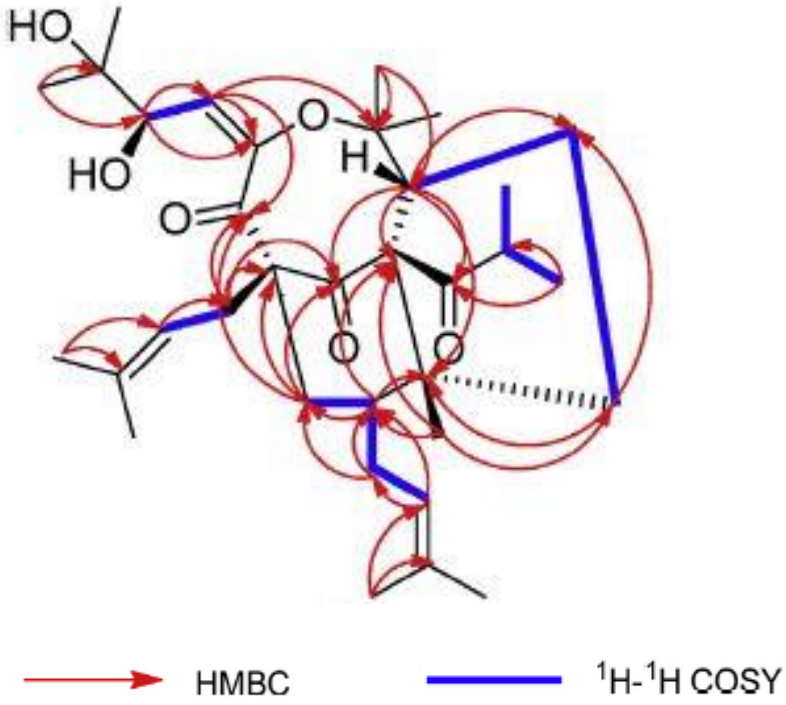
Compound 14, Hypericumoxide N, referred to as Liu14

***Phytochemistry* 2017, *142*, 38-50**

THE STARTING POINT



14



THE STARTING POINT

No.	14		
	δ_C	δ_H muti (J/Hz)	HMBC
1	206.3		
2	74.5		
3	58.5		
4	39.2	1.70 m	C-2,C-3,C-5,C-6, C-14,C-15, C-22
5	35.6	(Ha)2.61dd(14.1,3.0) 1.21 m (Hb)1.21 m	C-1, C-3, C-4, C-6,C-7, C-21
6	61.3		
7	195.8		
8	150.6		
9			
10	217.0		
11	42.7	3.20 dp(12.8, 6.4)	C-10, C-12, C-13
12	22.6	1.38 d (6.5)	C-10, C-11, C-13
13	19.8	1.12 d (6.4)	C-10, C-11, C-12
14	17.1	0.90 s	C-2, C-3, C-4, C-15
15	35.3	1.87ddd(12.6,9.7, 5.1), 1.42 m	C-2,C-3,C-4, C-14,C-16,C-17
16	25.7	2.32 m	C-3,C-15,C-17,C-18
17	55.3	2.88 t(10.0)	C-1,C-2,C-3,C-10,C-11, C-16,C-18,C-19,C-20
18	86.1		
19	24.9	1.19 s	C-17,C-18,C-20
20	24.2	1.15 s	C-17,C-18,C-19
21	29.4	2.02dd(12.8,5.5), 1.66 m	C-3,C-4,C-5, C-22, C-23
22	122.6	5.16 t(6.4)	C-4,C-21, C-23, C-24,C-25
23	133.4		
24	25.8	1.76 s	C-22, C-23, C-25
25	18.0	1.60 s	C-22, C-23, C-24
26	125.1	5.92 d(6.8)	C-7, C-8, C-18, C-27,C-28
27	74.0	4.19 d(6.8)	C-7,C-8,C-28, C-26,C-29,C-30
28	72.4		
29	26.6	1.19 s	C-27,C-28,C-30
30	25.8	1.15 s	C-27,C-28,C-29
31	34.1	2.35 m, 2.27 m	C-1,C-5,C-6, C-7,C-32,C-33
32	117.6	4.96 t(7.2)	C-6,C-31,C-33, C-34
33	135.2		
34	26.0	1.69 s	C-32,C-33,C-35
35	18.1	1.55 s	C-32,C-33,C-34

THE QUESTIONS

- Is structure **14** a logical consequence of the published data?
- Are there alternative **planar** structures?
  - If YES, is structure **14** the most likely one?

WAY 1 (NOT USED)

- Structure + Chemical shifts + Correlations + Assignments
  - *Everything is present in NMReDATA records*
  - Predict chemical shifts (and coupling constants) from Structure
    - Compare prediction and experiment according to assignment
  - Predict Correlations from Structure
    - Compare prediction and experiment according to assignment
  - Result: **Consistency of published data**

- Way 1 has to be implemented in the NMReDATA framework, so that authors and reviewers may check the consistency of published data.

WAY 2 (NOT USED)

- Structure + Chemical shifts + Correlations, **but no Assignments**
  - Predict chemical shifts (and coupling constants) from Structure
    - Match prediction and experiment, tentative assignment(s)
  - Predict Correlations from Structure
    - Match prediction and experiment, refined assignment(s)
  - Results **Consistency of Structure with Experimental data through the possibility of assigning resonances to atoms. Detection of assignment errors or imprecision.**

- *Way 2 may be implemented in the NMReDATA framework, so that authors and reviewers may check the quality of assignments.*

WAY 3

- Chemical shifts + Correlations, **but no Structure**
- **Computer-Assisted Structure Elucidation (CASE)**
  - Structure Elucidator (StrucEluc), ACD
  - CMC-se, Bruker
  - AssembleIt, ScienceSoft
  - COCON (J. Junker, M. Köck, T. Lindel, Frankfurt & Heidelberg)
  - SENECA (C. Steinbeck, Köln & Hinxton)
  - LSD (J.-M. Nuzillard, Reims)
- Result: Possible structures compatible with experimental data

- Authors may submit CASE input and output files to Journals as proofs of structure.

LIU14, LSD, INPUT DATA

- Compound Liu14 was used in July 2017 by Pr. Steinbeck for training students to the use of the LSD software in the context of a structural chemistry course.

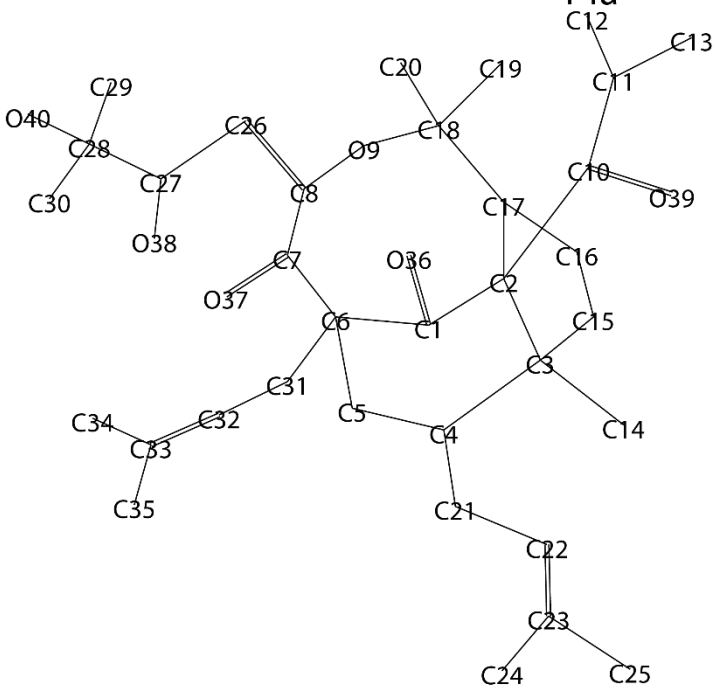
- LSD downloaded from www.univ-reims.fr/LSD
  - Free software, comes with installation and user's manuals
  - Tutorial article in *Magn. Reson. Chem.* **2017**, DOI: 10.1002/mrc.4612

LIU14, LSD, INPUT DATA

- Published list of atoms
  - From gross formula, $C_{34}H_{52}O_6$
  - 50 H bound to C, 2 OH groups, 3 C=O groups, 1 C-O-C group
  - Atom status from structure, consistent with ^{13}C NMR chemical shifts
- Published HSQC, HMBC and COSY data
  - Using published atom numbering, HSQC data is therefore implicit
  - Up to 3 ^{4}J HMBC correlations allowed
- Supplementary constraints:
  - Each sp^3 oxygen atom has all its neighbors in the list of carbon atoms at 55.3, 58.5, 61.3, 72.4, 74.0, 74.5, 86.1, and 150.6 ppm
  - Each carbon atom can carry one or less oxygen, but no more

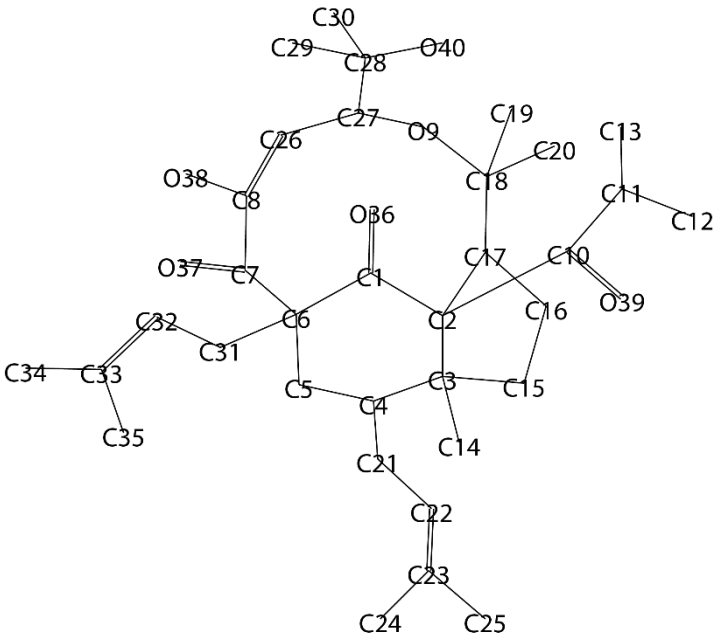
LIU14, 2 SOLUTIONS

14a



Published:
C8-09-C18 and C27-038H

14b



Unpublished:
C27-09-C18 and C8-O38H

SOLUTION COMPARISION

- Average absolute value of the differences of experimental and predicted values:
  - 14a: 3.0 ppm
  - 14b: 3.1 ppm
  - Using (an old) ACD predictor
  - Result is not significant
  - Other predictors are in favor of 14a

- No argument from nOe data

- 14b has an intra-cyclic conjugated enol function: Stability ?
- 14a has a conjugated enol ester function

- Other arguments???

WHAT'S NEXT?

- Possible conversion of NMReDATA files into CASE software input files
  - Automatic generation of status, chemical shift and correlation constraints
  - User adds supplementary constraints, if any

- Integration of tools for solution selection
  - Chemical shift predictors
  - Comparison with experimental data
  - Nice automatic structure drawing
  - Search for similarity/differences between structures

  - Handle 3D information…