

Article

# Sampling CASE Application for the Quality Control of Published Natural Product Structures

Lorena Martins Guimarães Moreira and Jochen Junker \* 

Oswaldo Cruz Foundation—CDTS, Rio de Janeiro 21040-900, Brazil; mglorena21@gmail.com

\* Correspondence: jochen.junker@fiocruz.br

**Abstract:** Structure elucidation with NMR correlation data is dicey, as there is no way to tell how ambiguous the data set is and how reliably it will define a constitution. Many different software tools for computer assisted structure elucidation (CASE) have become available over the past decades, all of which could ensure a better quality of the elucidation process, but their use is still not common. Since 2011, WEBCOCON has integrated the possibility to generate theoretical NMR correlation data, starting from an existing structural proposal, allowing this theoretical data then to be used for CASE. Now, WEBCOCON can also read the recently presented NMRreDATA format, allowing for uncomplicated access to CASE with experimental data. With these capabilities, WEBCOCON presents itself as an easily accessible Web-Tool for the quality control of proposed new natural products. Results of this application to several molecules from literature are shown and demonstrate how CASE can contribute to improve the reliability of Structure elucidation with NMR correlation data.

**Keywords:** theoretical NMR data; CASE; natural products; web-tool



**Citation:** Moreira, L.M.G.; Junker, J. Sampling CASE Application for the Quality Control of Published Natural Product Structures. *Molecules* **2021**, *26*, 7543. <https://doi.org/10.3390/molecules26247543>

Academic Editor: Antony J. Williams

Received: 22 May 2021

Accepted: 19 October 2021

Published: 13 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Software tools for computer assisted structure elucidation (CASE) of small molecules have been under development for over 50 years now. The methods used can roughly be divided into stochastic (S), deterministic (D), and hybrid (H), with hybrid representing various different combinations of stochastic and deterministic methods. The most prominent methods are fragment assemblers (H) [1–6], expert systems (H) [7–9], databases of  $^{13}\text{C}$  NMR chemical shifts or neural networks (S) [10–13], structure generation by reduction (S/H) [14], logic engines (D) [15], stochastic structure generators (S) [16], combinatorial brute force (D) [17–21], combinatorial brute force with restraints (D/H) [22,23], genetic algorithms (S) [24–26], simulated annealing (S) [27], convergent structure generation (S) [28,29], evolutionary algorithm (S) [30], fuzzy structure generation (S) [31], and expert systems with DFT (H) [32]. Altogether, 14 different methods were implemented in more than 20 different tools, some of which are freely usable, and many of them have not seen further development by their authors after the first publication. It can be observed that, out of the eight stochastic methods, only database methods are still actively developed, whereas both purely deterministic methods see ongoing development. However, the most active development happens for some of the hybrid methods, that combine stochastic and deterministic elements in various ways. Despite of all these efforts, none of these tools has become popular enough to be used for quality control of published structures by a wider community of researchers or journal editors. The  $^{13}\text{C}$  chemical shift-based tools have an outstanding position in this list, as they do not take NMR correlation data into account. They can either map chemical shifts to chemical neighborhoods, which in turn are assembled into a proposed constitution, or determine chemical neighborhoods for each carbon of a proposed constitution and predict a chemical shift. One such free to use tool is CSEARCH [10,13,33], which allows for an automated verification of  $^{13}\text{C}$  chemical shift assignments for a proposed structure combining a database and a neural network verification. This approach might fail if few similar structures are available, or if incorrect data

is contained in the database. For such applications, tools using NMR correlation data are better suited. One tool that uses NMR correlation data is COCON with the online interface WEBCOCON, which has been improved continuously over the past 20 years [22,34–39]. Today, COCON can be used freely via a web interface called WEBCOCON, that facilitates the input file creation, submission, and visualization of the results.

In a first step towards the use of CASE for quality control, the creation of theoretical NMR correlations from an existing constitution was implemented in WEBCOCON [37]. This way, the suggested structure is drawn on the web page, theoretical NMR correlations are created, and then they are submitted to a CASE run. It might seem simplistic, but, if WEBCOCON produces more than one structural proposal under these circumstances, it means that the constitution cannot be verified using NMR correlation data alone, and the use of other information is needed. One possibility is the inclusion of NOEs, which was explored for the first time in combination with genetic algorithms [26]. However, this approach was abandoned due to the very complex parametrization of the genetic algorithm, that defied automation [40]. Since then, NOEs have been used only after the constitution has been defined, in order to determine the stereochemistry of the resulting molecule [41,42]. WEBCOCON is the first CASE tool that can use NOEs in the ranking of the structural proposals, in a similar way as NOE data is frequently used in publications, but now backed up by structure calculations. It was already shown that, even with coarsely defined NOEs, the conformations obtained as results were similar to published X-ray structures. [39]. The exploration of  $^{13}\text{C}$  NMR chemical shift calculation is commonly accepted for the comparison of structural proposals, and it is carried out manually here. This is a first step towards a future quality control method for the structure discussion used in publications.

With the popularization of file formats, such as NMReDATA [43–45], the structure verification could be carried out automatically using experimental data. WEBCOCON can read NMReDATA files, thus offering the possibility of a structure discussion based on theoretical and experimental data.

New natural products published over the past years, whose structures were elucidated using NMR data were chosen from the literature and submitted to a structure discussion using WEBCOCON. As the molecules generally were not available as electronic files, structures were drawn in WEBCOCON. Theoretical NMR correlation data for  $^3J_{\text{HH-COSY}}$  and  $^{2,3}J_{\text{HC-HMBC}}$  correlations ( $^{2,3}J_{\text{HN-HMBC}}$  and 1,1-ADEQUATE correlations are generated but were not used) with carbon chemical shifts from an internal lookup table were created automatically for the structures, and the data was submitted to a CASE run. If more than one structure is proposed, the ranked results are compared automatically to the published structure. Additionally, the suggested alternative structures were compared manually to the experimental structures by back-calculating  $^{13}\text{C}$  chemical shifts using CS Chemdraw V15. When chemical shifts were used for the discussion, only the atoms with the largest expected deviations were selected, and the comparison was based on the average deviation observed ( $\bar{\Delta}$ ). Looking for further automation possibilities, the use of CSEARCH as an alternative to CS Chemdraw was investigated for one example, and, for two molecules, the use of the standard deviation ( $\sigma$ ) of the chemical shift deviations and their range was included. A standard deviation of 5.5 ppm is expected for the values back-calculated in CS Chemdraw, according to the manual.

$^3J_{\text{HH-COSY}}$  and  $^{2,3}J_{\text{HC-HMBC}}$  correlations are the most common experiments acquired that contain information about the heavy atom skeleton of a molecule. The three-bond-long COSY correlation between two hydrogen atoms indicates that the two bearing heavy atoms are directly connected. As both heavy atoms are required to be bonded to hydrogens, this experiment provides only limited information for proton poor compounds. The  $^{2,3}J_{\text{HC-HMBC}}$  correlations are observed between a hydrogen and a carbon that are two or three bonds away. This experiment requires only one hydrogen for the observation of the correlation; thus, it can be used to localize carbons without protons in the skeleton. However, as the observed correlations can originate over two or three bonds, and

differentiation normally is not possible, the incorporation of this data is more difficult. The use of CASE for the interpretation is recommended for this ambiguous data. Other experiments have been developed over the years to provide additional information about a molecule's skeleton, most prominently the  $^{2,3}J_{\text{HN-HMBC}}$ , 1,1-ADEQUATE, and H2BC. The  $^{2,3}J_{\text{HN-HMBC}}$  is similar to the carbon version described before but correlates to a nitrogen. The 1,1-ADEQUATE [22,34,35,46,47] and H2BC [48–50] are experiments that provide a  $^2J_{\text{HC-HMBC}}$  type of correlation and can be used complementary to the aforementioned  $^{2,3}J_{\text{HC-HMBC}}$  data. Theoretical  $^{2,3}J_{\text{HN-HMBC}}$  and 1,1-ADEQUATE correlations were not used in this analysis because this data is not routinely acquired and, therefore, also not included into publications.

WEBCOCON is a web-based interface to the program COCON. Initially, WEBCOCON helps in the creation of the input file for COCON, offering different functionalities for this purpose. With this input file, COCON creates a list of all constitutions that are compatible with the provided correlation data. These constitutions are ranked on the server by different methods and then transferred back to the web interface for visualization. Alternatively, the suggested constitutions can be downloaded in a SDF file and inspected offline by the user.

## 2. Results

The results resumed in Table 1 are just samples from the literature. Further automation in the ranking process will streamline the analysis for many more molecules.

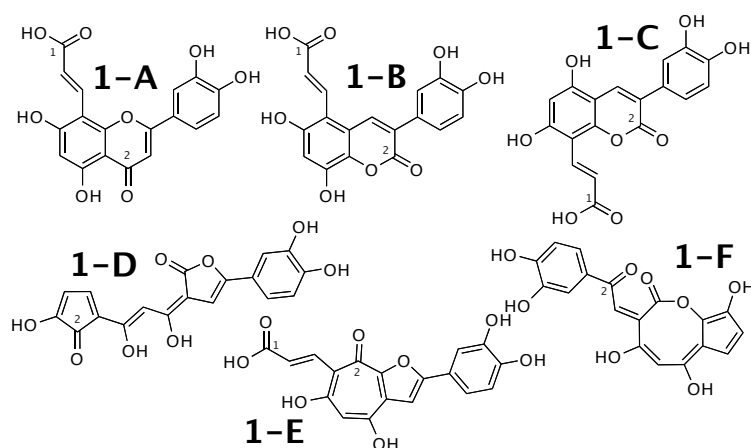
**Table 1.** Overview of the results.

Molecule	# Inspected	Confirmed
Luteolin 8-C-E-propenoic acid	6	yes
Tomentodiplacone	2	yes
Kadangustin A	5	no
Berkeleyamide D	2	no
14-norpseurotin A	2	no
Feruloylpodospermic acid A	2	no
Cochinchistemoninone	3	no
Milicifoline B	3	pending further analysis
5 $\alpha$ -cyprinol sulfate	4	yes
viridiol aka TAEMC161	>3	yes

### 2.1. Luteolin 8-C-E-Propenoic Acid

Luteolin 8-C-E-propenoic acid [51] was published as molecule **1-A** from Figure 1. WEBCOCON suggested 25 constitutions with the theoretical NMR correlation data set, including the published molecule. Out of those, only 6 underwent further inspection, as the others contained very restrained ring systems. The force field total energies calculated strongly favor molecules **1-A** and **1-B**.

The distinction of **1-A** from the other suggested constitutions can be done using  $^{13}\text{C}$  chemical shifts of selected atoms, when comparing experimental data with back-calculated data in Table 2. The back-calculated  $^{13}\text{C}$  chemical shift values for all carbons of all verified constitutions are in Figure A1.



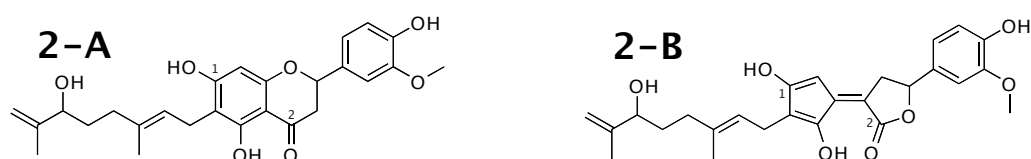
**Figure 1.** Luteolin 8-C-E-propenoic acid (**1-A**) and the 5 alternative constitutions suggested by WEBCOCON.

**Table 2.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for luteolin 8-C-E-propenoic acid (**1-A**) and the 5 suggested alternative constitutions.  $^{13}\text{C}$  chemical shifts were back-calculated for the atoms numbered in the molecules.  $\bar{\Delta}$  is the average of the deviations of the calculated to the experimental values.  $E_{\text{total}}$  is the total force field energy calculated in kcal/mol for the constitutions.

Atom	$\delta^{13}\text{C}$ [ppm]						
	Exp.	1-A	1-B	1-C	1-D	1-E	1-F
1	169.0	171.5	171.5	171.5	-	170.4	-
2	182.2	182.1	160.5	160.5	187.0	175.0	189.7
$\bar{\Delta}$		1.3	12.1	12.1	2.4	4.3	3.8
$E_{\text{total}}$ [kcal/mol]		26	29	45	80	74	303

## 2.2. Tomentodiplacone

Tomentodiplacone [52] was published as molecule **2-A** from Figure 2. WEBCOCON suggested 2 constitutions with the theoretical NMR correlation data set, including the published molecule. The force field total energy calculated for both molecules is almost equal.



**Figure 2.** Tomentodiplacone (**2-A**) and the alternative constitution suggested by WEBCOCON.  $^{13}\text{C}$  chemical shifts were back-calculated for the atoms numbered in the molecules.

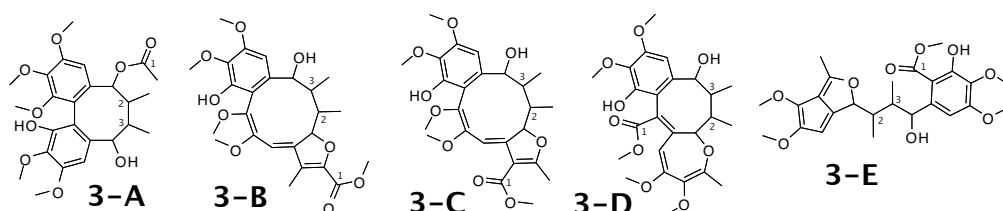
The distinction of **2-A** from the other suggested constitution can be done using  $^{13}\text{C}$  chemical shifts of selected atoms. The average deviation of the back-calculated chemical shifts from the experimental values is considerably smaller for **2-A**, as shown in Table 3. The back-calculated  $^{13}\text{C}$  chemical shift values for all carbons of all verified constitutions are in Figure A2.

**Table 3.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for tomentodiplacone (**2-A**) and the suggested alternative constitution.  $\bar{\Delta}$  is the average of the deviations of the calculated to the experimental values.  $E_{\text{total}}$  is the total force field energy calculated in kcal/mol for the constitutions.

Atom	$\delta^{13}\text{C}$ [ppm]		
	Exp.	2-A	2-B
1	164.4	162.5	182.6
2	196.4	196.8	170.0
$\bar{\Delta}$		1.2	22.3
$E_{\text{total}}$ [kcal/mol]		62	64

### 2.3. Kadangustin A

Kadangustin A [53] was published as molecule **3-A** from Figure 3. WEBCOCON suggested 13 constitutions with the theoretical NMR correlation data set, 5 of which are shown here, including the published molecule. The force field total energy calculated for these 5 molecules is very similar.



**Figure 3.** Kadangustin A (**3-A**) and the 4 alternative constitutions suggested by WEBCOCON.  $^{13}\text{C}$  chemical shifts were back-calculated for the atoms numbered in the molecules.

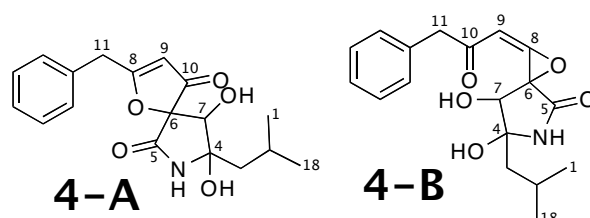
The distinction of **3-A** from the other suggested constitutions is possible using  $^{13}\text{C}$  chemical shifts of some selected atoms, when comparing experimental data with back-calculated data in Table 4 using the average deviation ( $\bar{\Delta}$ ). However, the standard deviation ( $\sigma$ ) and range ( $[\Delta_{\text{min}}^{\text{max}}]$ ) clearly favor **3-D** as solution, in spite of the higher total energy. This case should be verified in more detail. The back-calculated  $^{13}\text{C}$  chemical shift values for all carbons of all verified constitutions are in Figure A3.

**Table 4.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for kadangustin A (**3-A**) and the 4 suggested alternative constitutions.  $\bar{\Delta}$  is the average deviation,  $\sigma$  is the standard deviation of the chemical shifts, and  $[\Delta_{\text{min}}^{\text{max}}]$  is the range of the deviations of the calculated chemical shifts to the experimental values.  $E_{\text{total}}$  is the total force field energy calculated in kcal/mol for the constitutions.

Atom	$\delta^{13}\text{C}$ [ppm]					
	Exp.	3-A	3-B	3-C	3-D	3-E
1	172.0	170.2	162.1	165.0	165.4	171.4
2	39.3	31.3	33.6	33.6	35.9	46.7
3	34.8	34.8	38.7	38.7	31.1	41.8
$\bar{\Delta}$		3.3	6.5	5.5	4.6	5.0
$\sigma$		4.2	7.1	6.0	1.8	4.5
$[\Delta_{\text{min}}^{\text{max}}]$		8.0	13.8	10.9	6.6	8.0
$E_{\text{total}}$ [kcal/mol]		160	165	160	168	156

#### 2.4. Berkeleyamide D

Berkeleyamide D [54] was published as molecule **4-A** from Figure 4. WEBCOCON suggested 2 constitutions with the theoretical NMR correlation data set, including the published molecule. The force field total energy calculated for both molecules clearly favors **4-B**.



**Figure 4.** Berkeleyamide D (**4-A**) and the alternative constitution suggested by WEBCOCON.  $^{13}\text{C}$  and  $^1\text{H}$  chemical shifts were back-calculated for the atoms numbered in the molecules.

The confirmation of **4-A** as correct solution is also not possible using  $^{13}\text{C}$  and  $^1\text{H}$  chemical shift values of selected atoms, as the average deviation is smaller for **4-B**, as shown in Table 5. The back-calculated  $^{13}\text{C}$  and  $^1\text{H}$  chemical shift values using CS Chemdraw for all atoms of both verified constitutions are in Figures A4 and A5, respectively.

**Table 5.** Comparison of experimental and back-calculated  $^{13}\text{C}$  and  $^1\text{H}$  chemical shifts for berkeleyamide D (**4-A**) and the suggested alternative constitution.  $\bar{\Delta}$  is the average of the deviations of the calculated to the experimental values.  $E_{\text{total}}$  is the total force field energy calculated in kcal/mol for the constitutions. Deviations  $\gg 10$  ppm are labeled with “\*”.

Atom	Exp.	$\delta^{13}\text{C}$ [ppm]		Exp.	$\delta^1\text{H}$ [ppm]	
		4-A	4-B		4-A	4-B
4	84.9	76.6	87.2			
6	95.3	100.1	90.2			
7				4.41	4.42	4.0
8	197.7	165.5 *	198.0			
9	104.4	105.6	98.8	5.35	4.85	5.1
10	199.4	204.9	196.5			
11	37.4	45.0	55.7 *	3.98	3.22	4.2
$\bar{\Delta}$		8.5	4.9		0.5	0.3
$E_{\text{total}}$ [kcal/mol]		67	24			

Although the average deviations shown in Table 5 indicate that **4-B** matches the experimental chemical shift values better, for both suggested constitutions, there is one carbon with a chemical shift deviation  $\gg 10$  ppm (carbon 8 for **4-A** and carbon 11 for **4-B**). Therefore, all  $^{13}\text{C}$  chemical shift values back-calculated by CS Chemdraw (**M-I**) and CSEARCH (**M-II**) for all carbon atoms were included in a comparison in Table 6.

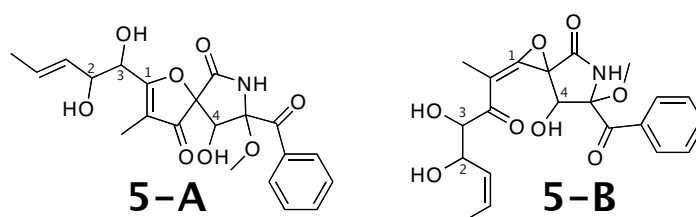
**Table 6.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for all carbons in berkeleyamide D (**4-A**) and the suggested alternative constitution, using CS Chemdraw (**M-I**) and CSEARCH (**M-II**).  $\sigma$  is the standard deviation and  $\left[ \begin{smallmatrix} \Delta_{max} \\ \Delta_{min} \end{smallmatrix} \right]$  is the range of the deviations between the back-calculated and experimental value.

Atom	Exp.	4-A		4-B	
		M-I	M-II	M-I	M-II
1	23.9	23.5	23.8	23.5	23.8
2	24.0	23.6	24.0	23.9	24.0
3	45.5	47.1	45.5	44.3	45.4
4	84.9	76.6	84.9	87.2	85.5
5	164.1	169.2	171.6	169.2	170.2
6	95.3	100.1	95.3	90.2	65.4
7	75.1	76.5	75.1	78.5	75.7
8	197.8	204.9	181.2	198.0	156.9
9	104.4	105.6	104.4	98.8	89.6
10	199.4	165.5	199.4	196.5	198.4
11	37.4	45.0	43.0	55.7	47.5
12	133.2	137.4	133.2	135.6	134.6
13	129.2	129.0	129.2	129.6	129.6
14	129.2	128.6	129.2	129.2	128.9
15	127.0	125.7	127.1	127.6	127.1
16	129.2	128.6	129.2	129.2	128.9
17	129.2	129.0	129.2	129.6	129.6
18	23.8	23.5	23.8	23.5	23.8
$\sigma$		9.0	4.6	5.1	12.5
$\left[ \begin{smallmatrix} \Delta_{max} \\ \Delta_{min} \end{smallmatrix} \right]$		41.5	24.1	23.9	51.0

The standard deviation  $\sigma$  and range  $\left[ \begin{smallmatrix} \Delta_{max} \\ \Delta_{min} \end{smallmatrix} \right]$  observed in the comparison of the experimental and the different back-calculated chemical shift values using both methods lead to a contradictory result. Whereas method **M-I** favors **4-B**, method **M-II** favors **4-A** as correct constitution. However, for both cases, the range  $\left[ \begin{smallmatrix} \Delta_{max} \\ \Delta_{min} \end{smallmatrix} \right]$  is very high, indicating that neither one of the constitutions might be correct. An analysis with experimental correlation data would be the next step, followed by a new verification of back-calculated carbon chemical shift values for all alternatives. If the NMR data was available as NMRReDATA, this would be straightforward.

### 2.5. 14-Norpseurotin A

14-norpseurotin A [55] was published as molecule **5-A** from Figure 5. WEBCOCON suggested 13 constitutions with the theoretical NMR correlation data set, including the published molecule. The force field total energies calculated favors **5-B** over **5-A**; all other alternatives show considerably higher values.



**Figure 5.** 14-norpseurotin A (**5-A**) and the alternative constitution suggested by WEBCOCON.  $^{13}\text{C}$  and  $^1\text{H}$  chemical shifts were back-calculated for the atoms numbered in the molecules.

The distinction of **5-A** from **5-B** is also not possible using  $^{13}\text{C}$  chemical shifts, when comparing experimental data with back-calculated data for selected atoms, as can be seen

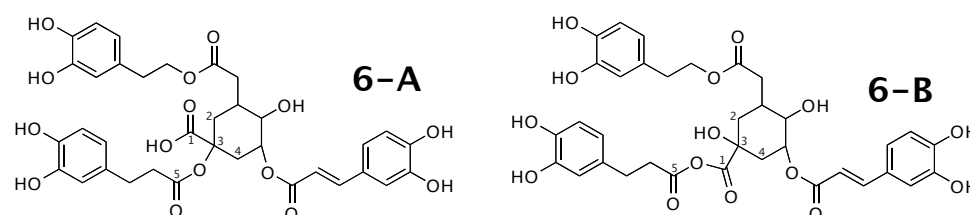
in Table 7. Actually, for **5-B** the back-calculated  $^{13}\text{C}$  and  $^1\text{H}$  chemical shifts match the experimental values better, as can be verified by the average deviation. This case should be verified in more detail. The back-calculated  $^{13}\text{C}$  and  $^1\text{H}$  chemical shift values for all verified constitutions are in Figures A6 and A7, respectively.

**Table 7.** Comparison of experimental and back-calculated  $^{13}\text{C}$  and  $^1\text{H}$  chemical shifts for 14-norpseurotin A (**5-A**) and the suggested alternative constitution.  $\bar{\Delta}$  is the average of the deviations of the calculated to the experimental values.  $E_{\text{total}}$  is the total force field energy calculated in kcal/mol for the constitutions.

Atom	$\delta^{13}\text{C}$ [ppm]			$\delta^1\text{H}$ [ppm]		
	Exp.	5-A	5-B	Exp.	5-A	5-B
1	185.7	167.3	188.0			
2	70.4	69.7	69.8	4.78	4.0	4.2
3	70.3	71.6	88.9	4.62	4.0	4.2
4	73.2	68.1	70.1	4.70	5.0	4.8
$\bar{\Delta}$		6.4	6.2		0.6	0.4
$E_{\text{total}}$ [kcal/mol]		160	148			

## 2.6. Feruloylpodospermic Acid A

Feruloylpodospermic acid A [56] was published as molecule **6-A** from Figure 6. WEB-COCON suggested 2 constitutions with the theoretical NMR correlation data set, including the published molecule. The force field total energy calculated for both molecules slightly favors **6-B**.



**Figure 6.** Feruloylpodospermic acid A (**6-A**) and the alternative constitution suggested by WEB-COCON.  $^{13}\text{C}$  chemical shifts were back-calculated for the atoms numbered in the molecules.

A comparison of the experimental and back-calculated  $^{13}\text{C}$  chemical shifts for the 5 carbons numbered in Figure 6 is shown in Table 8. The observed average deviation also slightly favors **6-B**. This case should be verified in more detail. The back-calculated  $^{13}\text{C}$  chemical shift values for all carbons of all verified constitutions are in Figure A8.

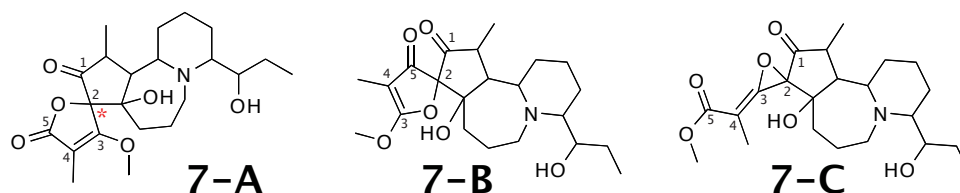
**Table 8.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for feruloylpodospermic acid A (**6-A**) and the suggested alternative constitution.  $\bar{\Delta}$  is the average of the deviations of the calculated to the experimental values.  $E_{\text{total}}$  is the total force field energy calculated in kcal/mol for the constitutions.

Atom	$\delta^{13}\text{C}$ [ppm]		
	Exp.	6-A	6-B
1	170.4	171.0	173.2
2	38.9	39.2	35.9
3	83.5	78.9	85.3
4	34.1	38.4	34.0
5	173.1	170.2	173.1
$\bar{\Delta}$		2.5	1.5
$E_{\text{total}}$ [kcal/mol]		75	56



### 2.7. Cochinchistemoninone

Cochinchistemoninone [57] was published as molecule **7-A** from Figure 7. WEBCOCON suggested 3 constitutions with the theoretical NMR correlation data set, including the published molecule. The force field total energy calculated for the molecules favors **7-C**.



**Figure 7.** Cochinchistemoninone (**7-A**) and the alternative constitution suggested by WEBCOCON.  $^{13}\text{C}$  chemical shifts were back-calculated for the atoms numbered in the molecules.

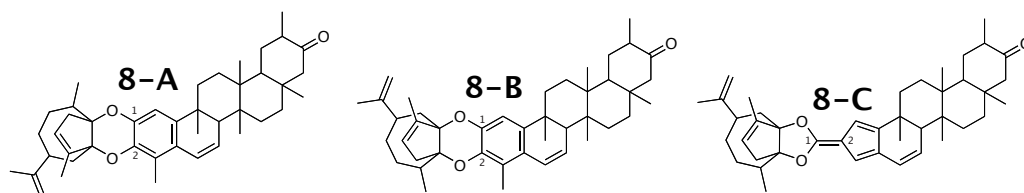
A comparison of the experimental and back-calculated  $^{13}\text{C}$  chemical shifts for the 5 carbons numbered in Figure 7 is shown in Table 9. The observed average deviation favors **7-A** as correct solution, but, in all suggested constitutions, the experimental value of the chemical shift of carbon 2 is considerably different from the back-calculated value. This case should be verified in more detail. The back-calculated  $^{13}\text{C}$  chemical shift values for all carbons of all verified constitutions are in Figure A9.

**Table 9.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for cochinchistemoninone (**7-A**) and the 2 suggested alternative constitutions.  $\bar{\Delta}$  is the average of the deviations of the calculated to the experimental values.  $E_{\text{total}}$  is the total force field energy calculated in kcal/mol for the constitutions.

Atom	Exp.	$\delta^{13}\text{C}$ [ppm]		
		7-A	7-B	7-C
1	208.9	211.2	211.3	211.3
2	90.6	117.6	127.4	103.3
3	169.2	171.8	182.6	191.4
4	98.4	98.7	88.1	95.6
5	173.1	175.2	197.9	167.3
$\bar{\Delta}$		6.9	17.5	9.2
$E_{\text{total}}$ [kcal/mol]		116	87	80

### 2.8. Milicifoline B

Milicifoline B [58] was published as molecule **8-A** from Figure 8. When using theoretical NMR correlation data from **8-A**, WEBCOCON suggested 3 constitutions, including the published molecule. The force field total energy calculated for the three molecules is very similar.



**Figure 8.** Milicifoline B (**8-A**) and the 2 alternative constitution suggested by WEBCOCON.

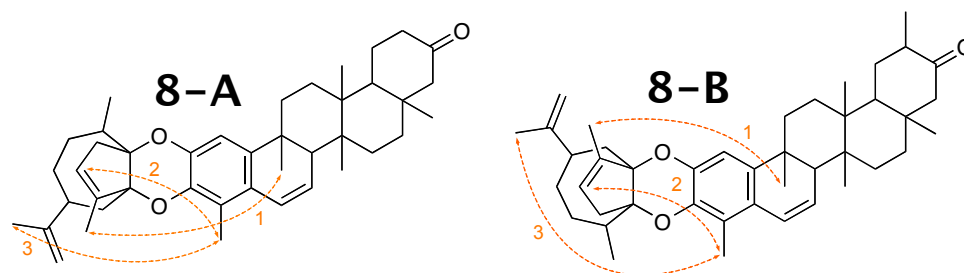
The experimental and back-calculated  $^{13}\text{C}$  chemical shifts for two carbons of all constitutions are shown in Table 10. Due to the large average deviation, the suggested constitution **8-C** can be excluded as possible solution. The average deviation favors **8-B**

over **8-A**, but without a strong differentiation. The back-calculated  $^{13}\text{C}$  chemical shift values for all carbons of all verified constitutions are in Figure A10.

**Table 10.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for milicifoline B (**8-A**) and the 2 suggested alternative constitutions.  $\bar{\Delta}$  is the average of the deviations of the calculated to the experimental values.  $E_{\text{total}}$  is the total force field energy calculated in kcal/mol for the constitutions.

Atom	$\delta^{13}\text{C}$ [ppm]			
	Exp.	8-A	8-B	8-C
1	144.1	145.5	145.5	139.6
2	142.8	136.9	144.4	91.7
$\bar{\Delta}$		3.6	1.5	27.8
$E_{\text{total}}$ [kcal/mol]		267	288	265

In their publication, the authors used 3 NOEs (see Figure 9) in order to decide between the two possible constitutions, so this information was also included into WEBCOCON.



**Figure 9.** NOEs (depicted in orange) used to identify the correct conformation of milicifoline B (**8-A**) in the literature and the corresponding contacts in **8-B**.

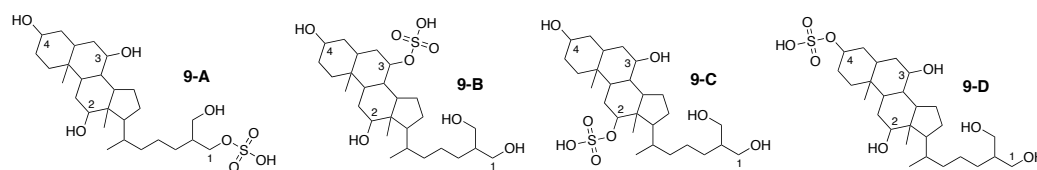
The distances measured after the WEBCOCON run are shown in Table 11, and they favor the originally published constitution. However, the distances measured in the calculated molecules are all  $> 500$  pm and, therefore, at the limit of measurable NOEs [59]. Due to this, no reliable decision between the two suggested constitutions can be made.

**Table 11.** Comparison of back-calculated distances for the NOEs for milicifoline B (**8-A**) and the suggested alternative constitution **8-B**.

NOE	$\Delta s$ [pm]	
	8-A	8-B
1	705	850
2	612	618
2	503	832

### 2.9. $5\alpha$ -Cyprinol Sulfate

$5\alpha$ -cyprinol sulfate [60] was published as molecule **9-A** from Figure 10. When using theoretical NMR correlation data from **9-A**, WEBCOCON suggested only one solution. However, for this molecule, experimental data is available in the NMRReDATA format. Thus, the WEBCOCON analysis was repeated using the published experimental data, resulting in 4 suggested constitutions, including the published molecule. The force field total energy calculated for the four molecules is very similar.



**Figure 10.** 5 $\alpha$ -cyprinol sulfate (**9-A**) and the 3 alternative constitutions suggested by WEBCOCON.

All constitutions have the same skeleton; they differ only in the positioning of the sulfate group. Hence,  $^{13}\text{C}$  chemical shifts are expected to vary mainly for the carbons adjacent to the sulfate group. The experimental and back-calculated  $^{13}\text{C}$  chemical shifts for these 4 carbons of all constitutions are shown in Table 12. The calculated average deviation over the 4 carbons for the suggested constitutions clearly favors the constitution **9-A**, matching the publication. The back-calculated  $^{13}\text{C}$  chemical shift values for all carbons of all verified constitutions are in Figure A11.

**Table 12.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for 5 $\alpha$ -cyprinol sulfate (**9-A**) and the 3 suggested alternative constitutions for selected carbons.  $\bar{\Delta}$  is the average of the deviations of the calculated to the experimental values.  $E_{\text{total}}$  is the total force field energy calculated in kcal/mol for the constitutions.

Atom	Exp.	$\delta^{13}\text{C}$ [ppm]			
		9-A	9-B	9-C	9-D
1	67.6	77.7	64.2	64.2	64.2
2	72.3	73.3	73.3	73.3	87.6
3	68.4	68.4	91.5	68.4	68.4
4	71.4	71.4	71.4	89.5	71.4
$\bar{\Delta}$		2.8	6.9	5.6	4.7
$E_{\text{total}}$ [kcal/mol]		42	48	56	53

### 2.10. Viridiol aka TAEMC161

Viridiol (**10-C**) was originally published in 1969, as derivative of viridin [61]. The crystal structure was published in 2013 [62], and, in recent years, the total synthesis has become available [63–66]. In 2000, TAEMC161 was isolated as metabolite from *Trichoderma hamatum*, and **10-B** was suggested as structure [67]. Shortly thereafter, the structure was revised as being identical to viridiol [68]. In 2010, the structure revision was reviewed by the use of CASE software, that confirmed previous findings [69]. The revisions were based mainly on the comparison of experimental and calculated carbon chemical shifts, using DFT and database methods.

The theoretical correlation data-based analysis of TAEMC161 reveals that this constitution could be described unambiguously by NMR correlation data. The same analysis with viridiol results in the 3 constitutions with similar total energies shown in Figure 11 and three more results with complex bridged ring systems that are not shown, with much higher total energies. From these 3 constitutions, **10-B** is the originally published metabolite TAEMC161, meaning that NMR correlation data cannot distinguish between viridiol and TAEMC161. Thus, other means of distinction need to be used. For this analysis, we have performed carbon chemical shift calculation for all atoms in all 3 constitutions using several methods: **M-I** was CSEARCH, **M-II** was CS Chemdraw, **M-IIIa** DFT with ORCA v5.0.1 using the conformation obtained by WEBCOCON and **M-IIIb** DFT with ORCA v5.0.1 using a DFT optimized conformation. The results are shown in Table 13 for **10-A**, Table 14 for TAEMC161 (**10-B**), and Table 15 for viridiol (**10-C**).

The resulting RMSD values for the chemical shift deviations observed for the different methods and constitutions show that the empirical methods outperform the DFT method in all conditions, as has already been observed previously [70,71]. In this case, **10-B** and

**10-C** cannot be distinguished by DFT, whereas both empirical methods clearly favor **10-C** as solution. Thus, the analysis confirms the published revisions but also proves that correlations data of viridiol are alone are not enough for a distinction of **10-B** and **10-C**.

**Table 13.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for all carbons in **10-A**. The largest deviations are bold in the table.

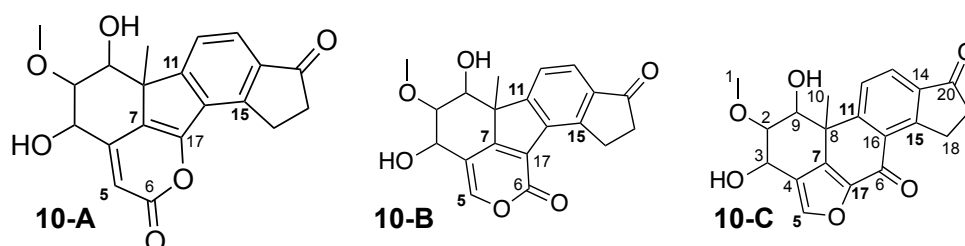
Atom	Exp.	M-I	M-II	M-IIIa	M-IIIb
1	60.7	60.8	57.8	61.0	60.4
2	81.7	82.0	88.9	93.6	89.5
3	61.6	70.4	66.1	68.8	69.3
4	122.1	138.8	164.0	154.2	158.8
5	<b>145.6</b>	<b>116.1</b>	<b>110.4</b>	<b>126.6</b>	<b>111.6</b>
6	173.4	161.1	162.6	174.0	165.8
7	<b>142.4</b>	<b>118.6</b>	<b>116.5</b>	<b>133.2</b>	<b>130.4</b>
8	42.3	47.3	44.8	55.5	55.1
9	71.7	71.9	80.7	80.0	75.9
10	30.5	28.0	21.0	22.7	22.1
11	<b>158.6</b>	157.2	<b>145.4</b>	<b>172.3</b>	<b>169.7</b>
12	127.3	125.4	125.1	131.4	128.8
13	127.3	126.6	123.7	131.6	130.9
14	129.8	133.9	133.9	142.8	144.9
15	158.0	150.7	137.8	154.4	158.5
16	136.9	131.5	124.5	136.9	137.6
17	<b>145.7</b>	146.4	148.6	<b>167.9</b>	<b>172.1</b>
18	28.4	29.0	24.3	28.8	28.6
19	36.5	35.7	36.6	38.3	39.3
20	206.7	206.8	206.2	213.0	210.5
RMSD		10.2	15.6	12.1	14.4

**Table 14.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for all carbons in TAEMC161 (**10-B**). The largest deviations are bold in the table.

Atom	Exp.	M-I	M-II	M-IIIa	M-IIIb
1	60.7	60.8	57.8	61.6	59.8
2	81.7	81.7	88.9	99.4	96.9
3	61.6	69.1	62.9	79.2	76.1
4	122.1	122.1	127.5	129.0	123.8
5	145.6	138.3	149.3	<b>158.5</b>	<b>158.4</b>
6	173.4	158.0	153.4	171.1	163.0
7	142.4	142.4	<b>153.0</b>	<b>165.6</b>	<b>174.0</b>
8	42.3	53.7	48.0	58.6	59.1
9	71.7	71.1	80.7	83.2	83.3
10	30.5	20.5	21.0	21.1	20.0
11	158.6	158.6	<b>145.4</b>	<b>171.9</b>	<b>166.3</b>
12	127.3	122.0	125.1	131.4	129.1
13	127.3	127.3	123.7	131.0	129.0
14	129.8	136.9	133.9	143.3	145.4
15	158.0	<b>139.2</b>	<b>137.8</b>	157.8	159.8
16	136.9	129.6	134.0	140.2	141.9
17	145.7	<b>123.5</b>	<b>124.5</b>	<b>138.0</b>	<b>129.5</b>
18	28.4	28.4	24.1	32.9	31.3
19	36.5	36.5	36.6	39.0	39.5
20	206.7	206.7	206.2	214.2	211.6
RMSD		8.8	9.8	11.0	12.0

**Table 15.** Comparison of experimental and back-calculated  $^{13}\text{C}$  chemical shifts for all carbons in viridiol (**10-C**). The largest deviations are bold in the table.

Atom	Exp.	M-I	M-II	M-IIIa	M-IIIb
1	60.7	60.8	57.5	59.8	59.5
2	81.7	82.0	78.9	94.5	90.2
3	61.6	68.5	64.7	63.8	63.8
4	122.1	121.0	125.9	130.9	128.5
5	145.6	146.3	144.5	155.5	153.2
6	173.4	178.9	173.0	153.8	154.8
7	142.4	140.4	146.6	149.6	149.4
8	42.3	45.7	39.2	51.9	49.0
9	71.7	73.7	85.9	81.7	76.2
10	30.5	20.2	23.5	31.5	32.4
11	158.6	153.9	149.3	<b>171.5</b>	<b>165.7</b>
12	127.3	127.3	125.3	137.4	132.1
13	127.3	127.4	128.4	133.3	132.2
14	129.8	133.2	134.1	141.4	143.9
15	158.0	<b>144.1</b>	<b>141.1</b>	165.0	168.8
16	136.9	126.5	126.4	141.2	137.7
17	145.7	148.0	145.3	<b>180.8</b>	<b>178.9</b>
18	28.4	28.4	23.5	32.9	33.2
19	36.5	36.2	36.2	38.9	39.6
20	206.7	206.3	206.2	214.2	211.7
RMSD		5.2	6.5	11.8	10.5

**Figure 11.** Viridiol (**10-C**), original TAEMC161 (**10-B**) and one other alternative constitution, all suggested by WEBCOCON. The positions with largest carbon chemical shift variations are marked in bold.

### 3. Discussion & Conclusions

CASE revealed more than one solution for all molecules discussed here. Since the suggested constitutions are compatible with the theoretical NMR correlation data set of the originally published structure, other means of identification of the correct constitution had to be explored. First, an inspection with “a chemists eye” of the results was carried out, supported by force field total energies, as WEBCOCON might have generated structures that are not likely to exist. In one case, the inclusion of NOE data from the original publication favors one of the possible solutions, but was not enough for a decision. Secondly, calculation of  $^{13}\text{C}$  NMR chemical shifts for the suggested constitutions was carried out and compared to the experimental data. However, for most examples, this was also not enough for a final decision.

The obtained results suggest that many of the new structures published using NMR correlation data could not be described unambiguously by theoretical NMR correlation data. Taking into account that the experimental data sets usually contain fewer data, even more new structures could be subjected to questions. Surveys of revised structures [32,69,71–73] clearly show that improvements as to how new small molecules are published are urgently needed. The use any of the software tools for CASE available, together with the popularization of more comprehensive data formats, such as NMRReDATA, will lead to an improved traceability of the structure elucidation process, as shown here for  $5\alpha$ -cyprinol sulfate. As it is, WEBCOCON already could easily be integrated into a quality control workflow.

**Author Contributions:** Conceptualization, L.M.G.M. and J.J.; methodology, J.J.; software, J.J.; investigation, L.M.G.M. and J.J.; writing, L.M.G.M. and J.J. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** All results shown in this article can be visualized by accessing the corresponding page on the WEBCOCON Server: [https://cocon-nmr.de/publication\\_data](https://cocon-nmr.de/publication_data) (accessed on 8 January 2021).

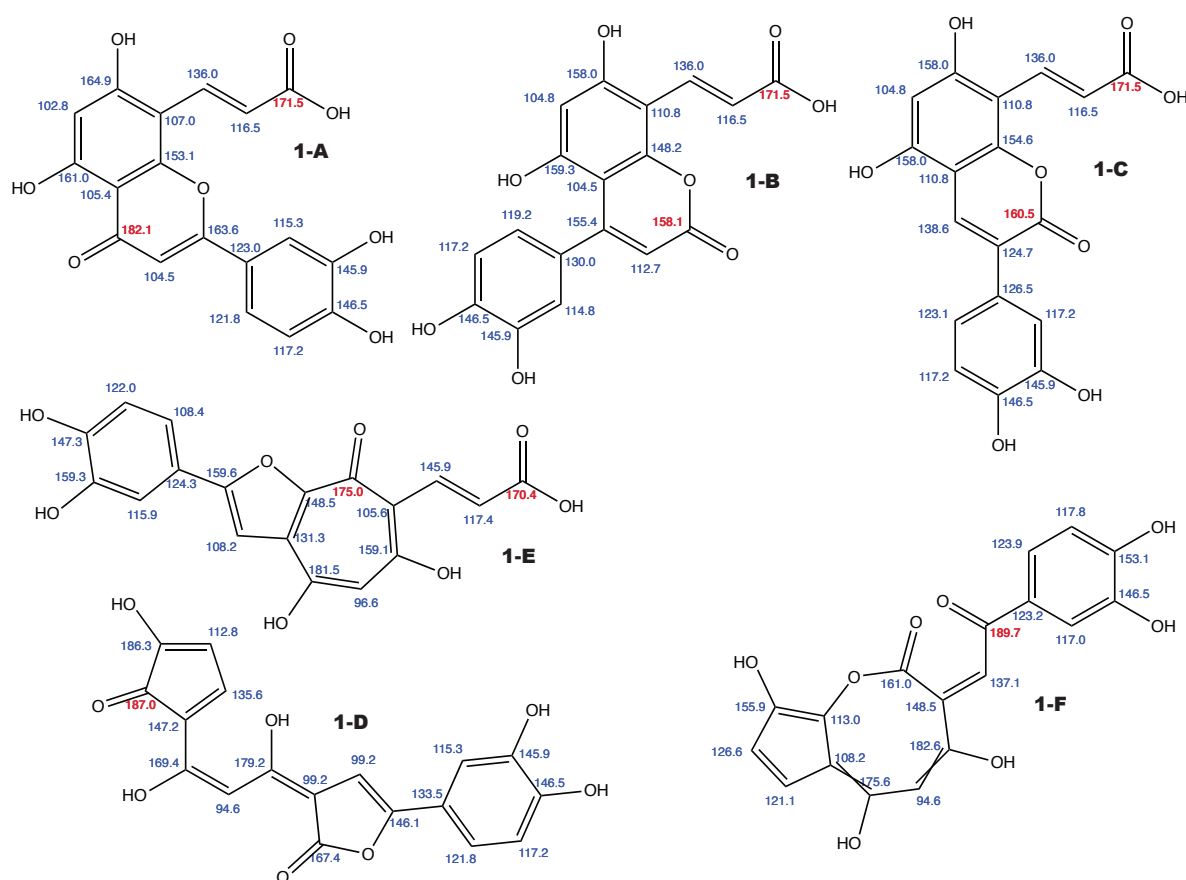
**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

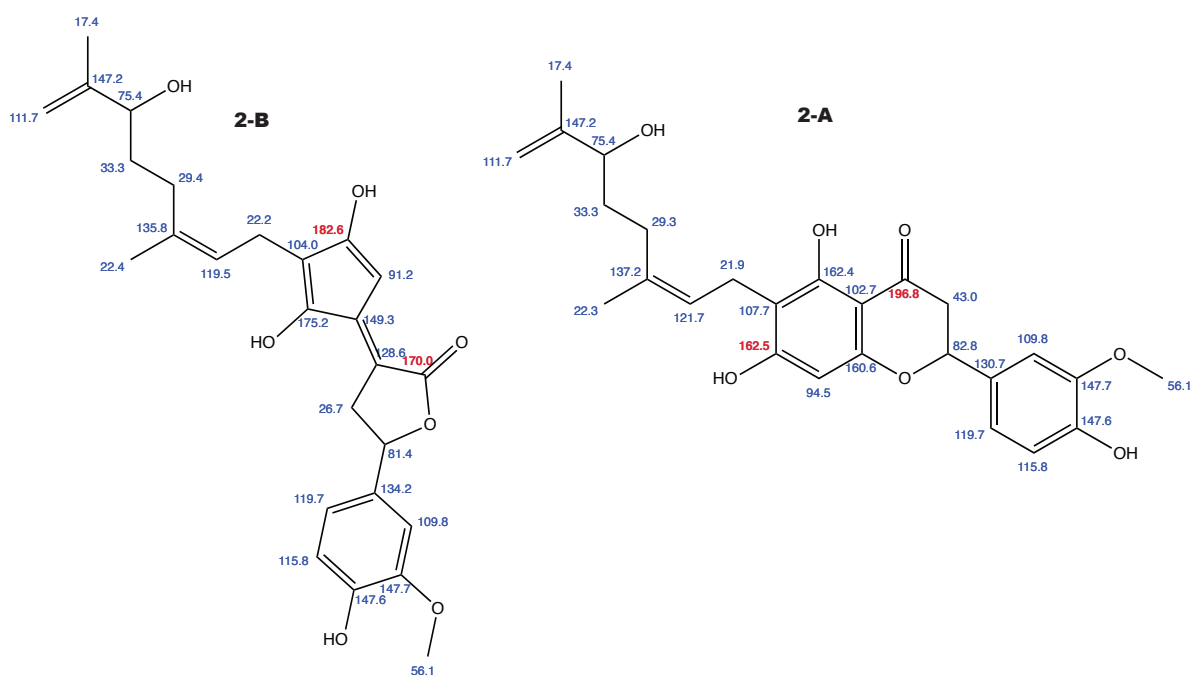
The following abbreviations are used in this manuscript:

ADEQ	1,1-ADEQUATE (“ $^2J_{CH}$ ” equivalent)
CASE	Computer Aided Structure Elucidation
COSY	$^1H,^1H$ -Correlated Spectroscopy ( $^3J_{HH}$ )
$\Delta s$	distance
$\delta^{13C}$	NMR carbon chemical shift
$\delta^1H$	NMR proton chemical shift
$\bar{\Delta}$	average deviation of the calculated chemical shift to the experimentally observed value
exp.	experimental
H2BC	HMQC-COSY
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
X-HMBC	$^1H,X$ -Heteronuclear Multiple Bond Correlation ( $^2J_{XH}$ and $^3J_{XH}$ )

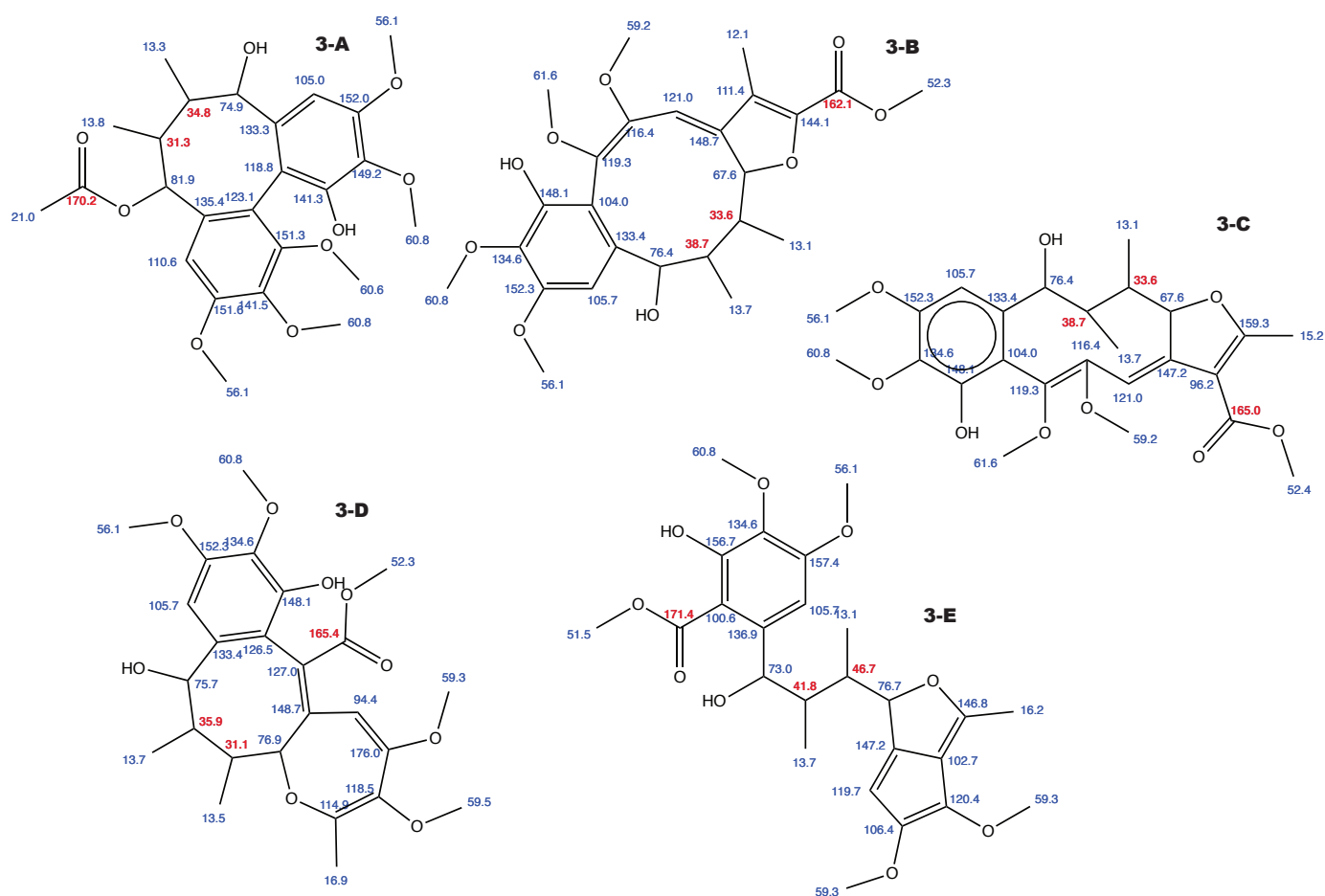
## Appendix A. All Assigned Back-Calculated Chemical Shifts



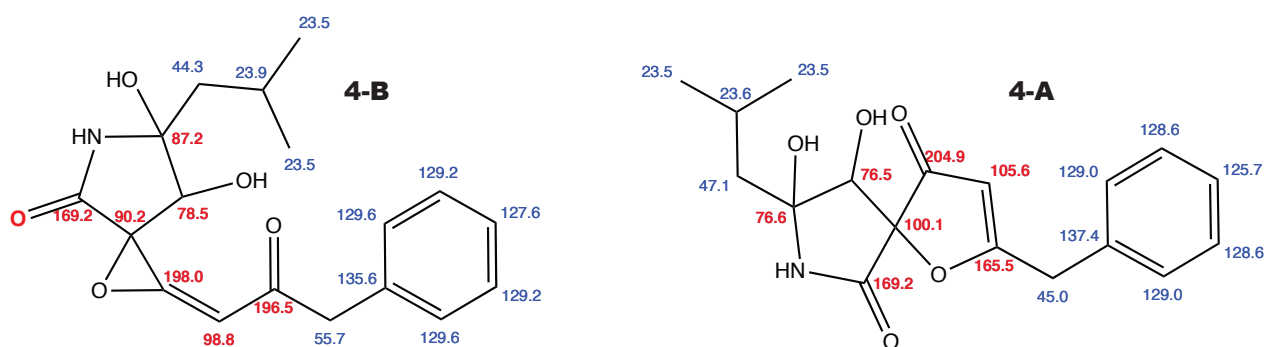
**Figure A1.** Luteolin 8-C-E-propenoic acid and the 5 alternative constitutions suggested by WEBCOCON with all back-calculated  $^{13}\text{C}$  chemical shifts. The  $^{13}\text{C}$  chemical shifts in red were used for the discussion.



**Figure A2.** Tomentodiplacone and the alternative constitution suggested by WEBCOCON with all back-calculated  $^{13}\text{C}$  chemical shifts. The  $^{13}\text{C}$  chemical shifts in red were used for the discussion.

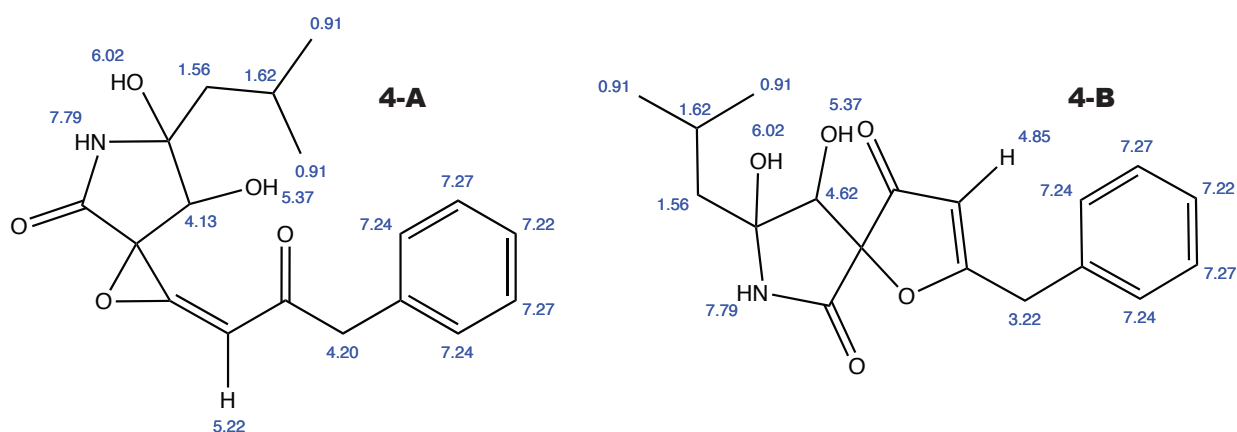


**Figure A3.** Kadangustin A and the 4 alternative constitutions suggested by WEBCOCON with all back-calculated  $^{13}\text{C}$  chemical shifts. The  $^{13}\text{C}$  chemical shifts in red were used for the discussion.

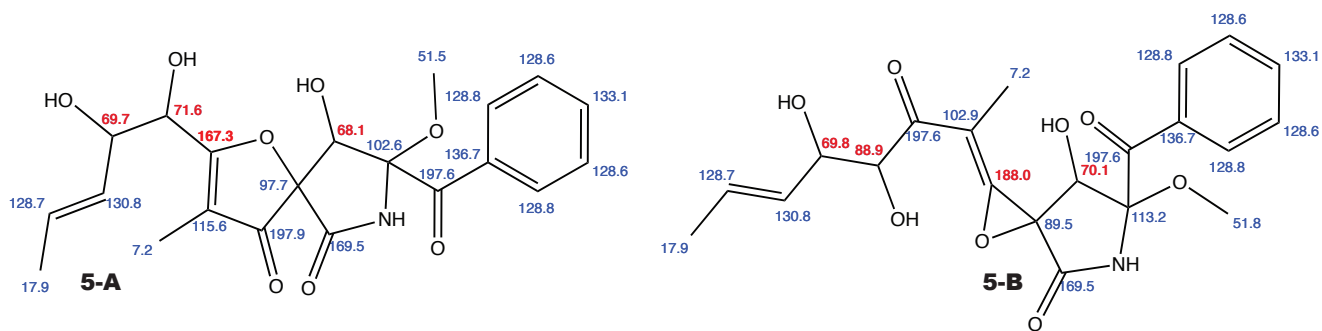


**Figure A4.** Berkeleyamide D and the alternative constitution suggested by WEBCOCON with all back-calculated  $^{13}\text{C}$  chemical shifts. The  $^{13}\text{C}$  chemical shifts in red were used for the discussion.

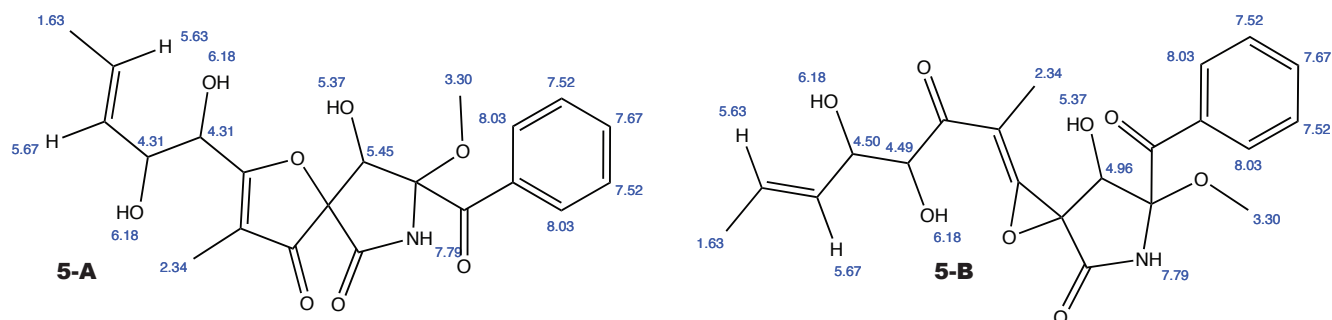




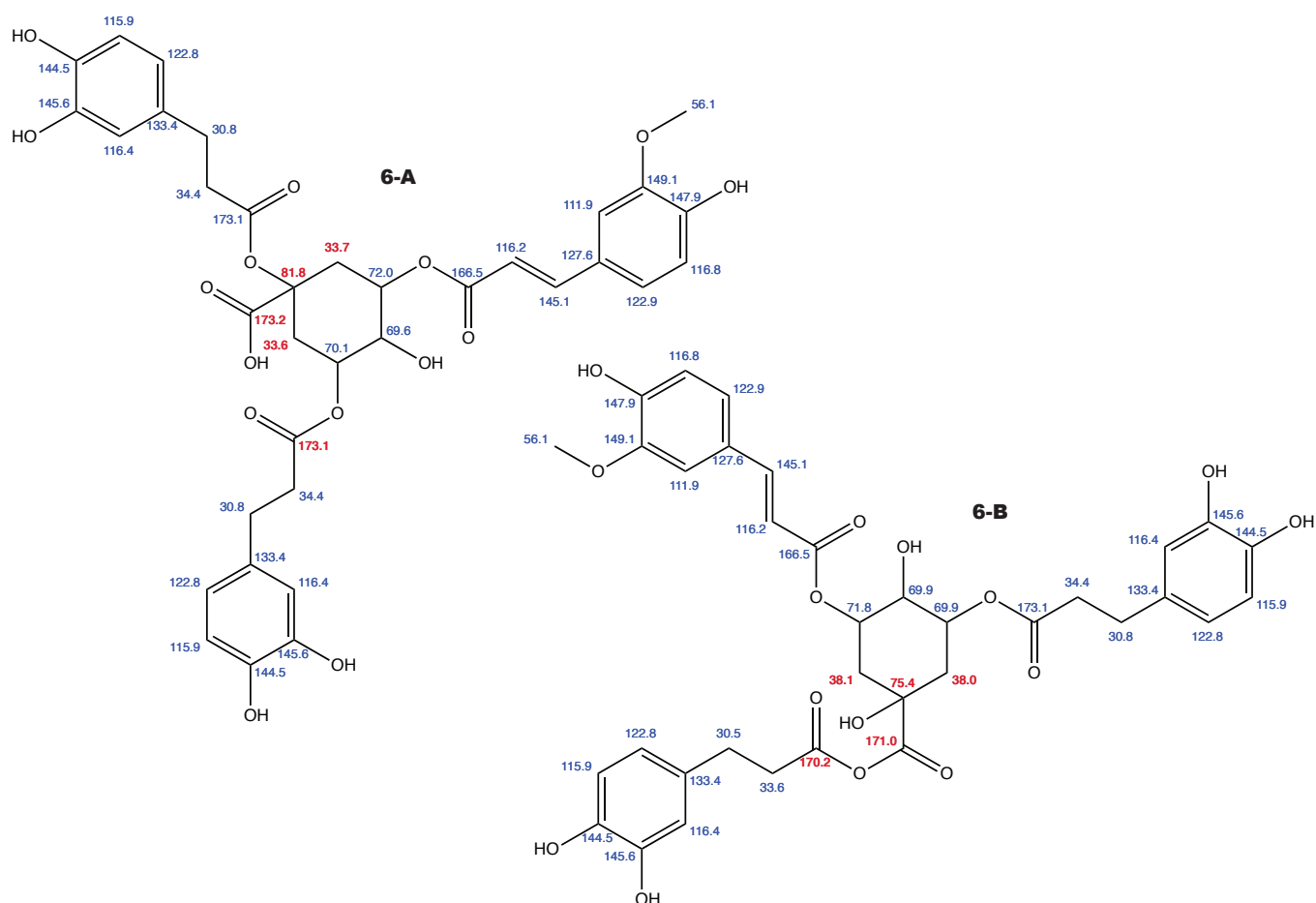
**Figure A5.** Berkeleyamide D and the alternative constitution suggested by WEBCOCON with all back-calculated  $^1\text{H}$  chemical shifts.



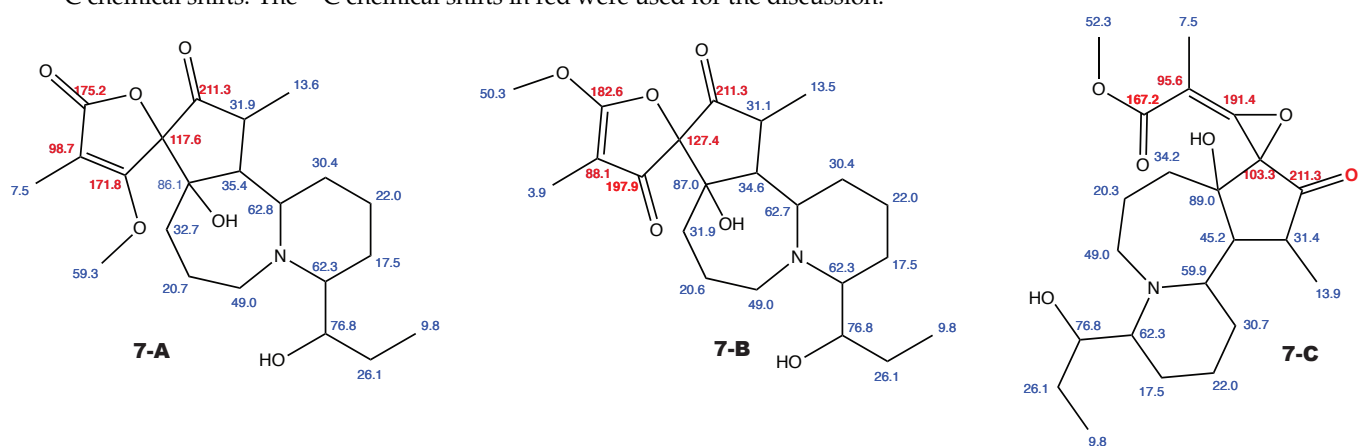
**Figure A6.** 14-norpseurotin A and the alternative constitution suggested by WEBCOCON with all back-calculated  $^{13}\text{C}$  chemical shifts. The  $^{13}\text{C}$  chemical shifts in red were used for the discussion.



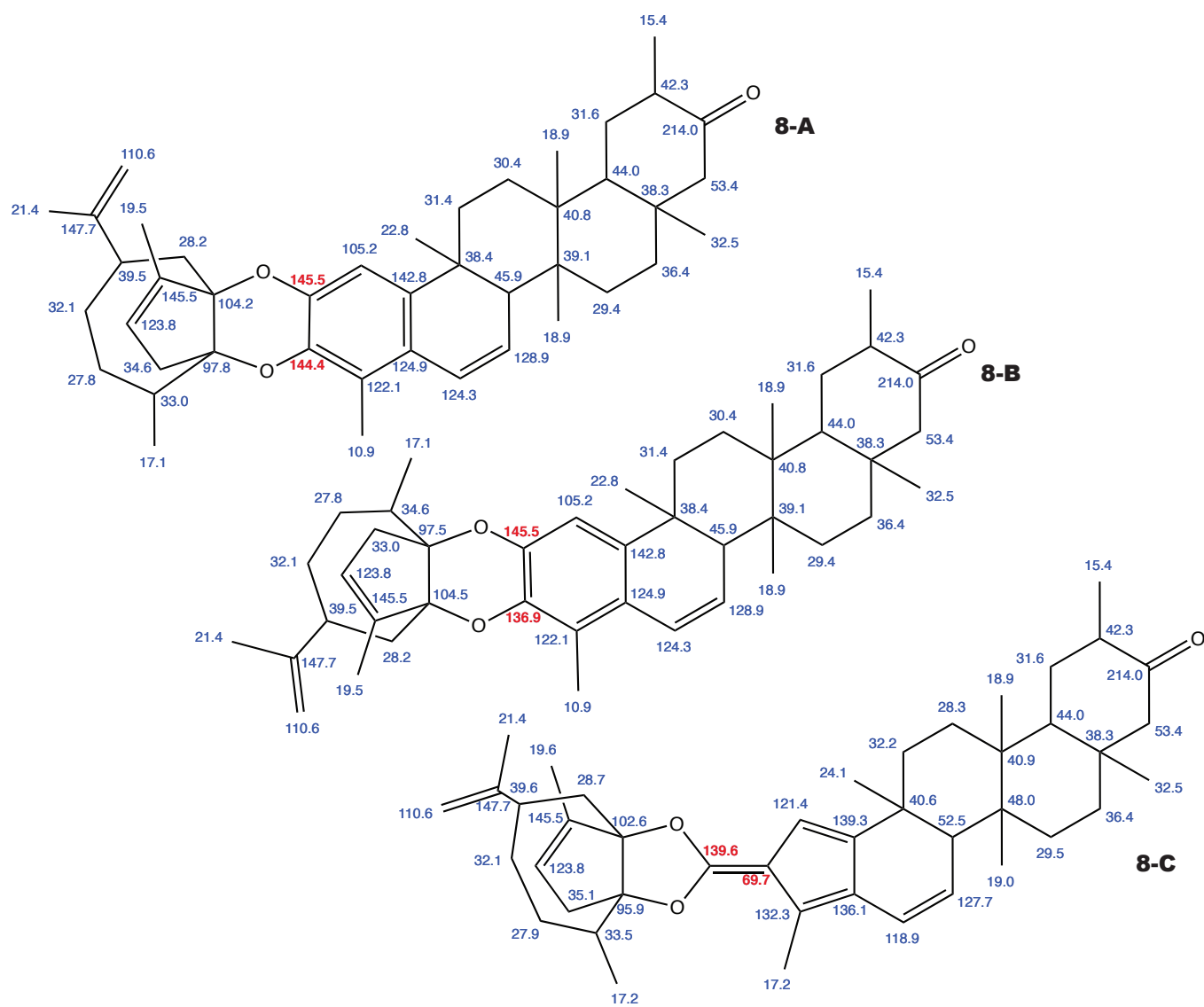
**Figure A7.** 14-norpseurotin A and the alternative constitution suggested by WEBCOCON with all back-calculated  $^1\text{H}$  chemical shifts.



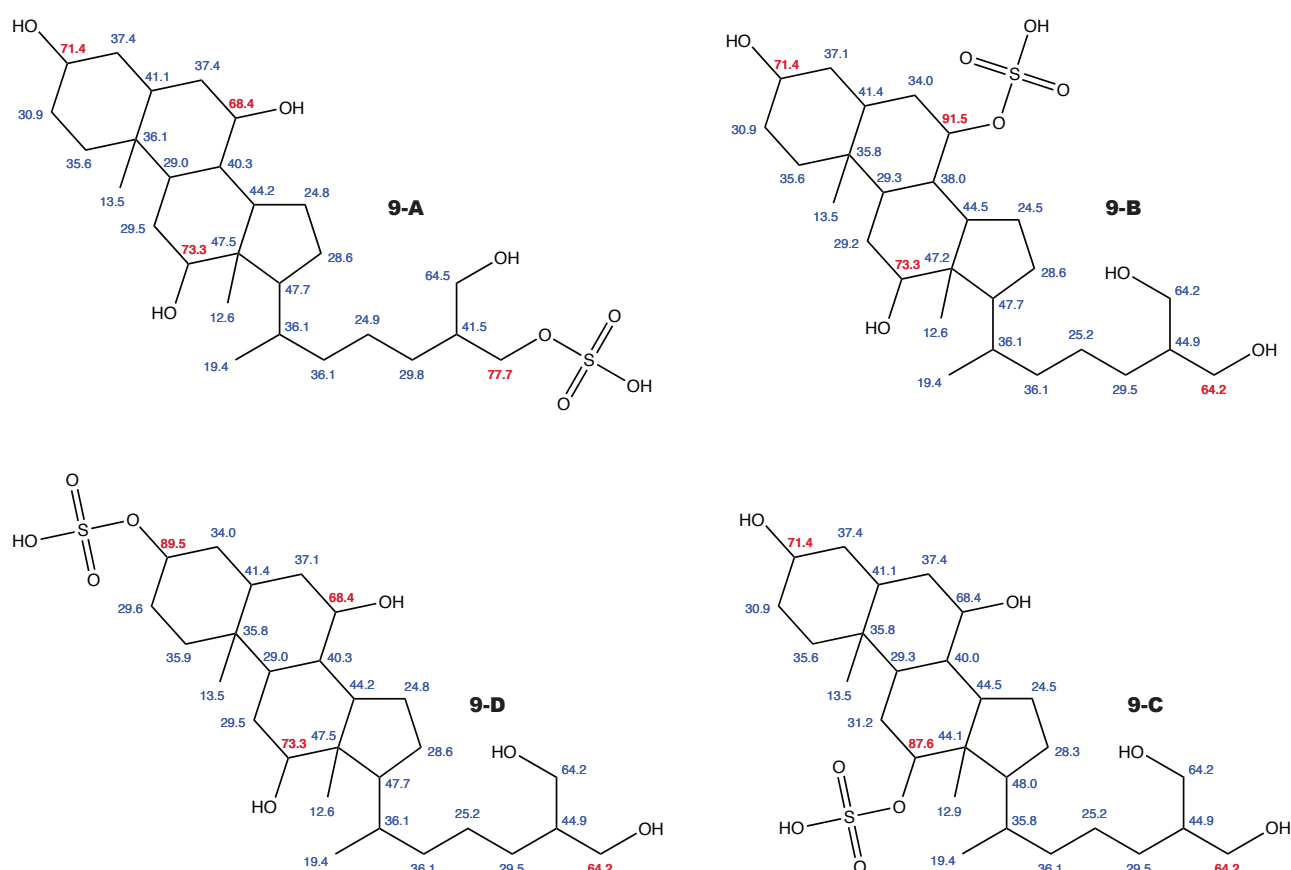
**Figure A8.** Feruloylpodospermic acid A and the alternative constitution suggested by WEBCOCON with all back-calculated  $^{13}\text{C}$  chemical shifts. The  $^{13}\text{C}$  chemical shifts in red were used for the discussion.



**Figure A9.** Cochinchistemoninone and the alternative constitution suggested by WEBCOCON with all back-calculated  $^{13}\text{C}$  chemical shifts. The  $^{13}\text{C}$  chemical shifts in red were used for the discussion.



**Figure A10.** Milicifoline B and the 2 alternative constitution suggested by WEBCOCON with all back-calculated  $^{13}\text{C}$  chemical shifts. The  $^{13}\text{C}$  chemical shifts in red were used for the discussion.



**Figure A11.** 5 $\alpha$ -cyprinol sulfate and the 3 alternative constitutions suggested by WEBCOCON with all back-calculated  $^{13}\text{C}$  chemical shifts. The  $^{13}\text{C}$  chemical shifts in red were used for the discussion.

## References

- Ichi Sasaki, S.; Kudo, Y.; Ochiai, S.; Abe, H. Automated chemical structure analysis of organic compounds: An attempt to structure determination by the use of NMR. *Mikrochim. Acta* **1971**, *59*, 726–742. [[CrossRef](#)]
- Yamasaki, T.; Abe, H.; Kudo, Y.; Sasaki, S.I. CHEMICS: A Computer Program System for Structure Elucidation of Organic Compounds. In *Computer-Assisted Structure Elucidation*; American Chemical Society: Washington, DC, USA, 1977; Chapter 8, pp. 108–125. [[CrossRef](#)]
- Sasaki, S.I.; Abe, H.; Hirota, Y.; Ishida, Y.; Kudo, Y.; Ochiai, S.; Saito, K.; Yamasaki, T. CHEMICS-F: A Computer Program System for Structure Elucidation of Organic Compounds. *J. Chem. Inf. Comput. Sci.* **1978**, *18*, 211–222. [[CrossRef](#)]
- Funatsu, K.; Sasaki, S.I. Recent advances in the automated structure elucidation system, CHEMICS. Utilization of two-dimensional NMR spectral information and development of peripheral functions for examination of candidates. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 190–204. [[CrossRef](#)]
- Zlatina, L.A.; Elyashberg, M.E. Generation and representation of stereoisomers of a molecular structure. *J. Struct. Chem.* **1992**, *32*, 528–533. [[CrossRef](#)]
- Pesek, M.; Juvan, A.; Jakoš, J.; Košmrlj, J.; Marolt, M.; Gazvoda, M. Database Independent Automated Structure Elucidation of Organic Molecules Based on IR,  $^1\text{H}$  NMR,  $^{13}\text{C}$  NMR, and MS Data. *J. Chem. Inf. Model.* **2021**, *61*, 756–763. [[CrossRef](#)] [[PubMed](#)]
- Gribov, L.A.; Elyashberg, M.E.; Raikhshtat, M.M. A new Approach to the Determination of Molecular Spatial Structures based on the use of Spectra and Computers. *J. Mol. Struct.* **1979**, *53*, 81–96. [[CrossRef](#)]
- Peng, C.; Yuan, S.; Zheng, C.; Hui, Y.; Wu, H.; Ma, K.; Han, X. Application of expert system CISOC-SES to the structure elucidation of complex natural products. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 814–819. [[CrossRef](#)]
- Elyashberg, M.E.; Blinov, K.A.; Williams, A.J.; Molodtsov, S.G.; Martin, G.E.; Martirosian, E.R. Structure elucidator: A versatile expert system for molecular structure elucidation from 1D and 2D NMR data and molecular fragments. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 771–792. [[CrossRef](#)] [[PubMed](#)]
- Kalchhauser, H.; Robien, W. CSEARCH: A Computer Program for Identification of Organic Compounds and Fully Automated Assignment of Carbon-13 Nuclear Magnetic Resonance Spectra. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 103–108. [[CrossRef](#)]
- Will, M.; Fachinger, W.; Richert, J.R. Fully automated structure elucidation - A spectroscopist's dream comes true. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 221–227. [[CrossRef](#)]
- Neudert, R.; Penk, M. Enhanced structure elucidation. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 244–248. [[CrossRef](#)]

13. Robien, W. The Advantage of Automatic Peer-Reviewing of <sup>13</sup>C-NMR Reference Data Using the CSEARCH-Protocol. *Molecules* **2021**, *26*, 3413. [[CrossRef](#)] [[PubMed](#)]
14. Christie, B.D.; Munk, M.E. Structure Generation by Reduction: A New Strategy for Computer-Assisted Structure Elucidation. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 87–93. [[CrossRef](#)] [[PubMed](#)]
15. Nuzillard, J.M.; Georges, M. Logic for structure determination. *Tetrahedron* **1991**, *47*, 3655–3664. [[CrossRef](#)]
16. Faulon, J.L. Stochastic Generator of Chemical Structure. 1. Application to the Structure Elucidation of Large Molecules. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1204–1218. [[CrossRef](#)]
17. Benecke, C.; Grund, R.; Hohberger, R.; Kerber, A.; Laue, R.; Wieland, T. MOLGEN+, a generator of connectivity isomers and stereoisomers for molecular structure elucidation. *Anal. Chim. Acta* **1995**, *314*, 141–147. [[CrossRef](#)]
18. Benecke, C.; Grüner, T.; Kerber, A.; Laue, R.; Wieland, T. Molecular structure generation with MOLGEN, new features and future developments. *Fresenius' J. Anal. Chem.* **1997**, *359*, 23–32. [[CrossRef](#)]
19. Meringer, M.; Schymanski, E.L. Small molecule identification with MOLGEN and mass spectrometry. *Metabolites* **2013**, *3*, 440–462. [[CrossRef](#)] [[PubMed](#)]
20. Gugisch, R.; Kerber, A.; Kohnert, A.; Laue, R.; Meringer, M.; Rücker, C.; Wassermann, A. MOLGEN 5.0, A Molecular Structure Generator. In *Advances in Mathematical Chemistry and Applications: Revised Edition*; Bentham Science Publishers: Sharjah, United Arab Emirates, 2015; Volume 1, Chapter 6, pp. 113–138. [[CrossRef](#)]
21. Kerber, A. MOLGEN, a generator for structural formulas. *Match-Commun. Math. Comput. Chem.* **2018**, *80*, 733–744.
22. Lindel, T.; Junker, J.; Köck, M. COCON: From NMR correlation data to molecular constitutions. *J. Mol. Model.* **1997**, *3*, 364–368. [[CrossRef](#)]
23. Badertscher, M.; Korytko, A.; Schulz, K.P.; Madison, M.; Munk, M.E.; Portmann, P.; Junghans, M.; Fontana, P.; Pretsch, E. Assemble 2.0: A structure generator. *Chemom. Intell. Lab. Syst.* **2000**, *51*, 73–79. [[CrossRef](#)]
24. Meiler, J.; Will, M. Automated Structure Elucidation of Organic Molecules from <sup>13</sup>C NMR Spectra Using Genetic Algorithms and Neural Networks. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1535–1546. [[CrossRef](#)]
25. Meiler, J.; Will, M. Genius: A genetic algorithm for automated structure elucidation from <sup>13</sup>C NMR spectra. *J. Am. Chem. Soc.* **2002**, *124*, 1868–1870. [[CrossRef](#)]
26. Smurnyy, Y.D.; Elyashberg, M.E.; Blinov, K.A.; Lefebvre, B.A.; Martin, G.E.; Williams, A.J. Computer-aided determination of relative stereochemistry and 3D models of complex organic molecules from 2D NMR spectra. *Tetrahedron* **2005**, *61*, 9980–9989. [[CrossRef](#)]
27. Steinbeck, C. SENECA: A Platform-Independent, Distributed, and Parallel System for Computer-Assisted Structure Elucidation in Organic Chemistry. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1500–1507. [[CrossRef](#)]
28. Korytko, A.; Schulz, K.P.; Madison, M.S.; Munk, M.E. HOUDINI: A New Approach to Computer-Based Structure Generation. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1434–1446. [[CrossRef](#)]
29. Schulz, K.P.; Korytko, A.; Munk, M.E. Applications of a HOUDINI-Based Structure Elucidation System. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1447–1456. [[CrossRef](#)]
30. Han, Y.; Steinbeck, C. Evolutionary-algorithm-based strategy for computer-assisted structure elucidation. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 489–498. [[CrossRef](#)] [[PubMed](#)]
31. Elyashberg, M.E.; Blinov, K.A.; Molodtsov, S.G.; Williams, A.J.; Martin, G.E. Fuzzy structure generation: A new efficient tool for Computer-Aided Structure Elucidation (CASE). *J. Chem. Inf. Model.* **2007**, *47*, 1053–1066. [[CrossRef](#)] [[PubMed](#)]
32. Elyashberg, M.; Blinov, K.; Williams, A. A systematic approach for the generation and verification of structural hypotheses. *Magn. Reson. Chem.* **2009**, *47*, 371–389. [[CrossRef](#)] [[PubMed](#)]
33. Robien, W. Computer-assisted peer reviewing of spectral data: The CSEARCH protocol. *Monatshfte Chem.* **2019**, *150*, 927–932. [[CrossRef](#)]
34. Köck, M.; Junker, J.; Maier, W.; Will, M.; Lindel, T. A COCON analysis of proton-poor heterocycles - Application of carbon chemical shift predictions for the evaluation of structural proposals. *Eur. J. Org. Chem.* **1999**, 579–586. [[CrossRef](#)]
35. Lindel, T.; Junker, J.; Köck, M. 2D-NMR-guided constitutional analysis of organic compounds employing the computer program COCON. *Eur. J. Org. Chem.* **1999**, 1999, 573–577. [[CrossRef](#)]
36. Junker, J.; Maier, W.; Lindel, T.; Köck, M. Computer-assisted constitutional assignment of large molecules: COCON analysis of Ascomycin. *Org. Lett.* **1999**, *1*, 737–740. [[CrossRef](#)]
37. Junker, J. Theoretical NMR correlations based structure discussion. *J. Cheminformatics* **2011**, *3*, 27. [[CrossRef](#)]
38. Junker, J. Statistical filtering for NMR based structure generation. *J. Cheminformatics* **2011**, *3*, 31. [[CrossRef](#)] [[PubMed](#)]
39. Köck, M.; Lindel, T.; Junker, J. Incorporation of 4J-HMBC and NOE Data into Computer-Assisted Structure Elucidation with WebCocon. *Molecules* **2021**, *26*, 4846. [[CrossRef](#)]
40. Elyashberg, M.E.; Argyropoulos, D. Author response to the manuscript of the article titled “Computer Assisted Structure Elucidation (CASE): Current and Future Perspectives”. *Magn. Reson. Chem.* **2020**. [[CrossRef](#)]
41. Elyashberg, M.; Argyropoulos, D. Computer Assisted Structure Elucidation (CASE): Current and future perspectives. *Magn. Reson. Chem.* **2021**, *59*, 669–690. [[CrossRef](#)]
42. Navarro-Vázquez, A.; Gil, R.R.; Blinov, K. Computer-Assisted 3D Structure Elucidation (CASE-3D) of Natural Products Combining Isotropic and Anisotropic NMR Parameters. *J. Nat. Prod.* **2018**, *81*, 203–210. [[CrossRef](#)]

43. Pupier, M.; Nuzillard, J.M.; Wist, J.; Schlörer, N.E.; Kuhn, S.; Erdelyi, M.; Steinbeck, C.; Williams, A.J.; Butts, C.; Claridge, T.D.; et al. NMRReDATA, a standard to report the NMR assignment and parameters of organic compounds. *Magn. Reson. Chem.* **2018**, *56*, 703–715. [[CrossRef](#)]
44. Trevorow, P.; Jeannerat, D. Reporting on the first NMRReDATA Symposium, 26 September 2019, Porto, Portugal. *Magn. Reson. Chem.* **2020**, *58*, 218–222. [[CrossRef](#)]
45. Kuhn, S.; Wieske, L.H.E.; Trevorow, P.; Schober, D.; Schlörer, N.E.; Nuzillard, J.; Kessler, P.; Junker, J.; Herráez, A.; Farès, C.; et al. NMRReDATA: Tools and applications. *Magn. Reson. Chem.* **2021**, *59*, 792–803. [[CrossRef](#)]
46. Reif, B.; Köck, M.; Kerssebaum, R.; Kang, H.; Fenical, W.; Griesinger, C. ADEQUATE, a New Set of Experiments to Determine the Constitution of Small Molecules at Natural Abundance. *J. Magn. Reson. Ser. A* **1996**, *118*, 282–285. [[CrossRef](#)]
47. Köck, M.; Junker, J.; Lindel, T. Impact of the H-1,N-15-HMBC experiment on the constitutional analysis of alkaloids. *Org. Lett.* **1999**, *1*, 2041–2044. [[CrossRef](#)]
48. Nyberg, N.T.; Duus, J.; Sørensen, O.W. Heteronuclear two-bond correlation: Suppressing heteronuclear three-bond or higher NMR correlations while enhancing two-bond correlations even for vanishing  $2J_{CH}$ . *J. Am. Chem. Soc.* **2005**, *127*, 6154–6155. [[CrossRef](#)] [[PubMed](#)]
49. Nyberg, N.T.; Duus, J.; Sørensen, O.W. Editing of H2BC NMR spectra. *Magn. Reson. Chem.* **2005**, *43*, 971–974. [[CrossRef](#)] [[PubMed](#)]
50. Petersen, B.O.; Vinogradov, E.; Kay, W.; Würtz, P.; Nyberg, N.T.; Duus, J.; Sørensen, O.W. H2BC: A new technique for NMR analysis of complex carbohydrates. *Carbohydr. Res.* **2006**, *341*, 550–556. [[CrossRef](#)] [[PubMed](#)]
51. Zhao, J.; Pawar, R.S.; Ali, Z.; Khan, I.A. Phytochemical investigation of *Turnera diffusa*. *J. Nat. Prod.* **2007**, *70*, 289–292. [[CrossRef](#)]
52. Šmejkal, K.; Chudík, S.; Klouček, P.; Marek, R.; Cvačka, J.; Urbanová, M.; Julínek, O.; Kokoška, L.; Šlapetová, T.; Holubová, P.; et al. Antibacterial C-geranylflavonoids from *Paulownia tomentosa* fruits. *J. Nat. Prod.* **2008**, *71*, 706–709. [[CrossRef](#)]
53. Gao, X.M.; Pu, J.X.; Huang, S.X.; Yang, L.M.; Huang, H.; Xiao, W.L.; Zheng, Y.T.; Sun, H.D. Lignans from *Kadsura angustifolia*. *J. Nat. Prod.* **2008**, *71*, 558–563. [[CrossRef](#)] [[PubMed](#)]
54. Stierle, A.A.; Stierle, D.B.; Patacini, B. The berkeleyamides, amides from the acid lake fungus *Penicillium rubrum*. *J. Nat. Prod.* **2008**, *71*, 856–860. [[CrossRef](#)]
55. Zhang, M.; Wang, W.L.; Fang, Y.C.; Zhu, T.J.; Gu, Q.Q.; Zhu, W.M. Cytotoxic alkaloids and antibiotic nordammarane triterpenoids from the marine-derived fungus *Aspergillus sydowi*. *J. Nat. Prod.* **2008**, *71*, 985–989. [[CrossRef](#)] [[PubMed](#)]
56. Tsevegsuren, N.; Edrada, R.; Lin, W.; Ebel, R.; Torre, C.; Ortlepp, S.; Wray, V.; Proksch, P. Biologically active natural products from Mongolian medicinal plants *Scorzonera divaricata* and *Scorzonera pseudodivaricata*. *J. Nat. Prod.* **2007**, *70*, 962–967. [[CrossRef](#)] [[PubMed](#)]
57. Wang, Y.Z.; Tang, C.P.; Dien, P.H.; Ye, Y. Alkaloids from the roots of *Stemona saxorum*. *J. Nat. Prod.* **2007**, *70*, 1356–1359. [[CrossRef](#)]
58. Gutiérrez, F.; Estévez-Braun, A.; Ravelo, Á.G.; Astudillo, L.; Zárate, R. Terpenoids from the medicinal plant *Maytenus ilicifolia*. *J. Nat. Prod.* **2007**, *70*, 1049–1052. [[CrossRef](#)] [[PubMed](#)]
59. Marius Clore, G. NMR in Structural and Cell Biology. In *Encyclopedia of Cell Biology*; Bradshaw, R.A., Stahl, P.D., Eds.; Academic Press: Waltham, MA, USA, 2016; pp. 98–107. [[CrossRef](#)]
60. Hahn, M.; von Elert, E.; Bigler, L.; Díaz Hernández, M.D.; Schloerer, N.E.  $5\alpha$ -Cyprinol sulfate: Complete NMR assignment and revision of earlier published data, including the submission of a computer-readable assignment in NMRReDATA format. *Magn. Reson. Chem.* **2018**, *56*, 1201–1207. [[CrossRef](#)] [[PubMed](#)]
61. Moffatt, J.S.; Bu'Lock, J.D.; Yuen, T.H. Viridiol, a steroid-like product from *Trichoderma viride*. *J. Chem. Soc. Chem. Commun.* **1969**, 839a. [[CrossRef](#)]
62. Andersson, P.F.; Broberg, A.; Lundberg, D. The furanosteroid viridiol. *Acta Crystallogr. Sect. Struct. Rep. Online* **2013**, *69*, o467–o468. [[CrossRef](#)]
63. Carreira, E.M.; Wolleb, H. Synthesis of (–)-Viridin and (–)-Viridiol. *Synfacts* **2017**, *13*, 0791. [[CrossRef](#)]
64. Del Bel, M.; Abela, A.R.; Ng, J.D.; Guerrero, C.A. Enantioselective Chemical Syntheses of the Furanosteroids (–)-Viridin and (–)-Viridiol. *J. Am. Chem. Soc.* **2017**, *139*, 6819–6822. [[CrossRef](#)] [[PubMed](#)]
65. Ji, Y.; Xin, Z.; He, H.; Gao, S. Total Synthesis of Viridin and Viridiol. *J. Am. Chem. Soc.* **2019**, *141*, 16208–16212. [[CrossRef](#)] [[PubMed](#)]
66. Carreira, E.M.; Wolleb, H. Synthesis of (–)-Viridin and (–)-Viridiol. *Synfacts* **2020**, *16*, 0119. [[CrossRef](#)]
67. Sakuno, E.; Yabe, K.; Hamasaki, T.; Nakajima, H. A new inhibitor of 5'-hydroxyaverantin dehydrogenase, an enzyme involved in aflatoxin biosynthesis, from *Trichoderma hamatum*. *J. Nat. Prod.* **2000**, *63*, 1677–1678. [[CrossRef](#)]
68. Wipf, P.; Kerekes, A.D. Structure reassignment of the fungal metabolite TAEMC161 as the phytotoxin viridiol. *J. Nat. Prod.* **2003**, *66*, 716–718. [[CrossRef](#)] [[PubMed](#)]
69. Elyashberg, M.; Williams, A.J.; Blinov, K. Structural revisions of natural products by Computer-Assisted Structure Elucidation (CASE) systems. *Nat. Prod. Rep.* **2010**, *27*, 1296–1328. [[CrossRef](#)]
70. Elyashberg, M.; Blinov, K.; Smurnyy, Y.; Churanova, T.; Williams, A. Empirical and DFT GIAO quantum-mechanical methods of  $^{13}\text{C}$  chemical shifts prediction: Competitors or collaborators? *Magn. Reson. Chem.* **2010**, *48*, 219–229. [[CrossRef](#)] [[PubMed](#)]
71. Marcarino, M.O.; Zanardi, M.M.; Sarotti, A.M. The Risks of Automation: A Study on DFT Energy Miscalculations and Its Consequences in NMR-based Structural Elucidation. *Org. Lett.* **2020**, *22*, 3561–3565. [[CrossRef](#)] [[PubMed](#)]

- 
72. Nicolaou, K.C.; Snyder, S.A. Chasing molecules that were never there: Misassigned natural products and the role of chemical synthesis in modern structure elucidation. *Angew. Chem.-Int. Ed.* **2005**, *44*, 1012–1044. [[CrossRef](#)]
  73. Elyashberg, M.; Blinov, K.; Molodtsov, S.; Williams, A. Elucidating 'undecipherable' chemical structures using computer-assisted structure elucidation approaches. *Magn. Reson. Chem.* **2012**, *50*, 22–27. [[CrossRef](#)]