A Platinum Open Access Journal for Organic Chemistry

Paper

Free to Authors and Readers

DOAJ Seal

Arkivoc **2022**, part iii, 0-0

to be inserted by editorial office

On the role of ”mental leaps” in small-molecule structure elucidation by NMR spectroscopy

Márton Weber,[a,b] Áron Szigetvári,[a] Mónika Halmai,[b] Péter Keglevich,[b] and Csaba Szántay Jr*[a]

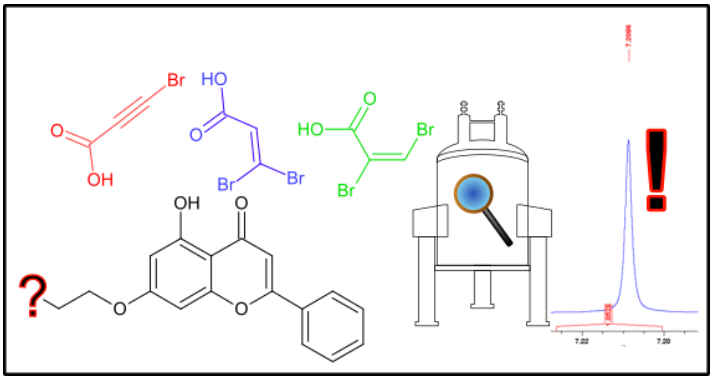
[a]*Spectroscopic Research Department, Gedeon Richter Plc., Gyömrői út 19-21, Budapest H-1103, Hungary*
[b]*Department of Organic Chemistry and Technology, Budapest University of Technology and Economics, Szt. Gellért tér 4, Budapest, H-1111, Hungary*
Email: cs.szantay@richter.hu

Received mm-dd-yyyy **Accepted Manuscript** mm-dd-yyyy **Published on line** mm-dd-yyyy

Dates to be inserted by editorial office

Abstract

While scientific publications often mention uncertainties or mistakes associated with a scientific conclusion due to technical limitations, they very rarely note even the slightest possibility of a man-made error in connection with a research program. However, such human mistakes are ubiquitous in science, and usually come from what we call “mental traps” or a lack of professional experience. In this article, we outline a simple model of the complex learning process through which the professional expertise of a scientist evolves. Key elements of this development are the “mental leaps”, the moments when one redefines one’s approach to new problems. In this rather unusual paper, we report on resolving three intriguing NMR-based structure-elucidation problems involving this learning process by an NMR spectroscopist employed in an industrial R&D & QC environment.



Keywords: NMR Spectroscopy, learning cycle, vindoline, chrysine, scientific thinking, mental traps, COSY, HSQC, HMBC, ROESY, 1,1-ADEQUATE

Introduction

This article is unusual in that its aims are to go beyond being merely technical. Overlaying our scientific discussion is a strong intention to assert and demonstrate that psychological factors are inherently and inevitably intertwined with scientific thinking, and with the learning process of a scientist. This approach, as well as the need to expand upon such ideas, is partly based on, and driven by, our previous work along similar lines.[1-5] Within that general mindset, it is our intention to demonstrate that the widespread myth according to which powerful modern NMR methods relegate the structure elucidation of small molecules to an almost mechanical and fool-proof process is not valid in practice. A profound understanding of this fact, and that of the underlying cognitive psychological reasons, are absolutely crucial when aiming for high-quality and high-confidence structure determination – especially so in the kind of high-stakes, high-pressure pharmaceutical industrial research environment that the present authors represent. Moreover, such understanding is imperative for the proper assessment of the scope and (inherent) limitations of NMR structure verification and structure-elucidation software which have, by now, become an almost compulsory requisite of the NMR toolkit.[1,5] Yet, these psychological aspects keep being overlooked and unappreciated by the scientific community (intriguingly, this ignorance also has its deep-seated psychological causes).

Our starting point is the tenet that, in science, the majority of seminal works come from scientists who have not only gained sufficient technical knowledge, but have undergone the intellectually challenging or even grinding process of solving unexpected, unfamiliar problems, and who have endured the mental hardships of making mistakes, of pursuing false hypotheses, etc. During this learning process, one has to grapple not only with intellectual and experimental difficulties, but with the innate fallibilities of human thinking which affect even the most astute minds in the form of "mental traps".[1] Recognizing and overcoming these traps is a particularly difficult task, and gaining true competence in their control is almost impossible without having fallen into some of them.

For the sake of the present discussion, it will be useful to envisage the aforementioned learning process – in a much simplified form – as a "learning cycle" (Figure 1).

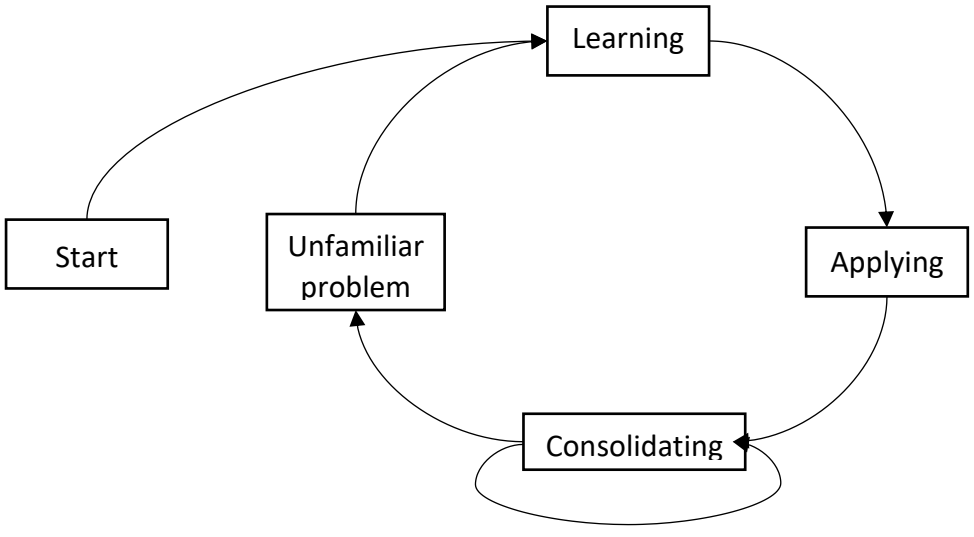


Figure 1: The "Learning Cycle".

 ©**AUTHOR(S)**

According to this very simple model, we first enter the cycle by acquiring some basic knowledge in the pertinent field of science – this may be called the (initial) "learning phase". Then, we often put our knowledge into practice by starting to apply it on real-life problems that can be solved within the boundaries of this knowledge. Next, by repetitively encountering problems whose solution requires no more than this knowledge base, we gain robust practical experience. This "consolidatory phase", in which we become able to solve familiar problems quickly and correctly, can be highly gratifying since it generates a sense of competence and develops self-confidence. Almost inevitably, however, we will at some point come across an unfamiliar problem that calls for competence extending beyond our comfort zone of knowledge, so we are forced out of this cozy mindset. We experience mental stress, whereby we often treat the novel problem as if it was a familiar one. Furthermore, sometimes it is difficult to recognize the true nature of these unfamiliar problems. Such problems can delay or completely stop our self-improvement and may be hard to overcome. We will demonstrate these aspects of the learning cycle through some examples below.

New problems can be solved through the application of new approaches. This requires critical and creative thinking as well as novel knowledge built upon our existing solid-knowledge base. But since we are reflexively guided by our brain to think in the same ways as what have given us all those gratifications in the consolidatory phase, new-solution schemes are often difficult to find. A "mental leap" is required that leads us to gain new insights, a new mindset, and to discover new methods that can be used if a similar problem occurs in the future. Taking such a mental leap can be hard. In that regard the catalytic role of peers – especially those who have already experienced such jumps – is often crucially important. Once the leap has been made, this becomes the new learning phase, ultimately leading us to a new application phase, followed by a more advanced consolidatory phase of a higher quality and efficiency.

With the above ideas in mind, this paper aims to demonstrate our notion of the learning process via three interesting and instructive NMR-based structure elucidation problems that we have recently encountered in our own work. Although two of us (Á.S. and C.S.) have extensive experience in the methodology and art of NMR structure determination, one of us (M.W.) is a relative newcomer to the field. Therefore, we conduct our discussion below in the spirit of reflecting on his mental evolution during the learning cycle. (For the sake of argumentation convenience, throughout our discussion we use a first-person plural narrative voice; in essence, collectively identifying with M.W.) Nevertheless, we attest that the ideas conveyed below through these three examples should be illuminating not only for any novice NMR spectroscopist, but also for the more seasoned ones (as was the case for Á.S. and C.S.). Indeed, growing via the learning cycle is a never-ending story since even very experienced spectroscopists can encounter, for example, a new family of compounds whose successful structure determination may prove to be an outside-of-comfort-zone challenge.

Before discussing these examples, we first need to make the following note. Our spectroscopic research facility operates principally on a holistic structure-elucidation philosophy, meaning that, in order to achieve high-level structural confidence and spectral characterization, all analytes submitted for structural analysis are thoroughly and almost simultaneously investigated by NMR and MS (and IR if needed).[2] The interpretations are harmonized amongst these techniques before the end result is disclosed to the client. Occasionally, however, either the MS or the NMR measurements may lag behind the other due to technical or other difficulties, in which case either the MS or the NMR investigation will move ahead on its own due to clientele time pressure. The MS or NMR interpretation is then relayed to the client under the mutual understanding that the final result (which may potentially overrule the temporary result) will be released when both measurements will have been interpreted, and any possible disagreements reconciled. The following examples represent such cases in which the MS results caught up with the ongoing NMR investigations only at a later stage due to technical reasons.

 ©AUTHOR(S)

Results and Discussion

To demonstrate the aforementioned learning process in practice, we describe the structure elucidation of three selected examples among newly synthesized vindoline (Figure 2, compound **1**) and chrysin (Figure 2, compound **2**) derivatives. *Vinca* alkaloids, such as vindoline, can be found in the leaves of *Catharanthus roseus,* and several members of them exhibit antitumor properties.[6] The development of novel anticancer agents exhibiting better pharmacological profiles is a constant need in drug discovery. To that end, a promising approach is the combination of easily accessible pharmacophores (such as compounds **1** and **2**) into a single molecule (**3**) via a suitable linker.[7]

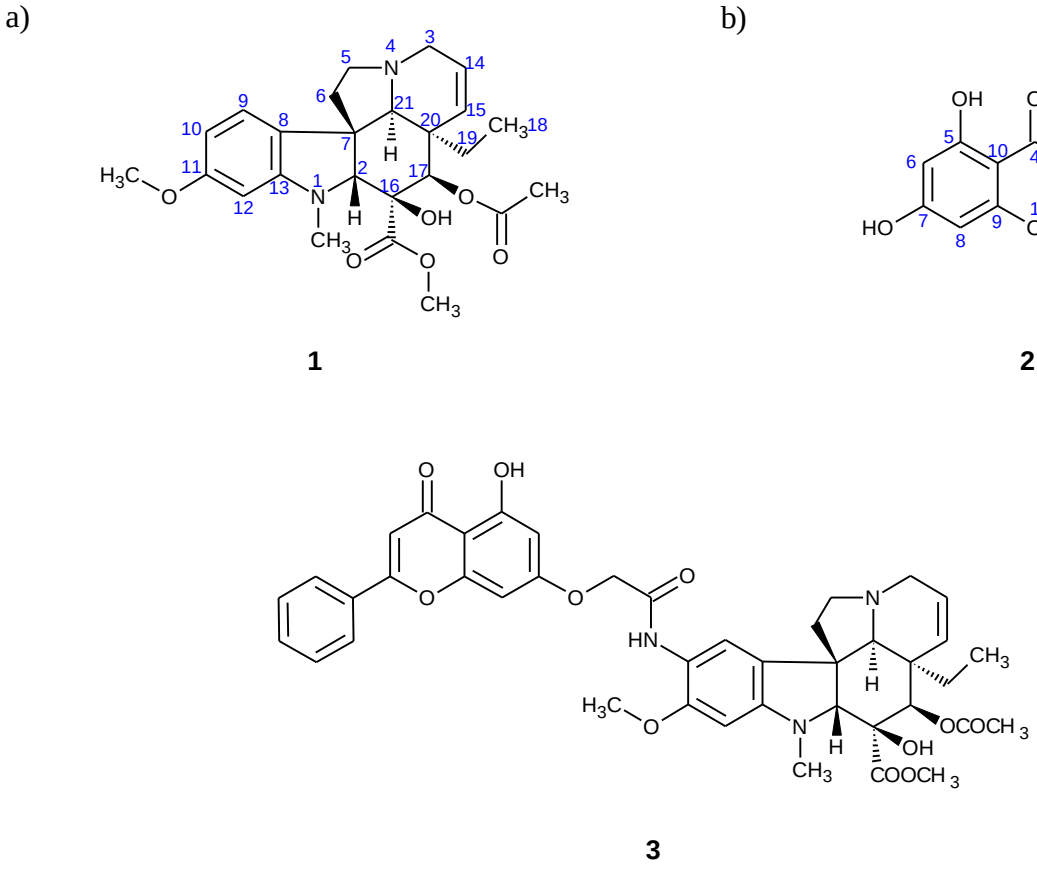


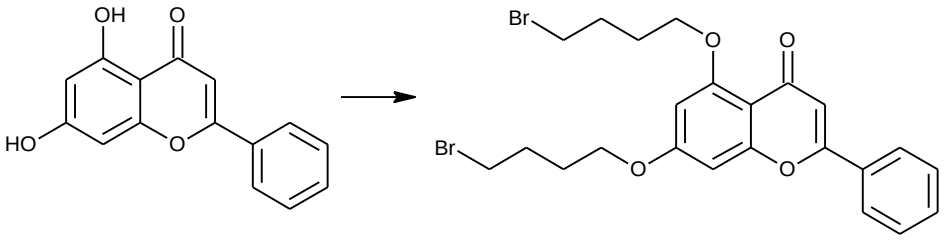
Figure 2: The structure of vindoline (**1**), chrysin (**2**), and a vindoline–chrysin hybrid (**3**)

Example 1

The first example discusses the structure elucidation of an unexpected product formed by allowing an alkylating agent to react with chrysin; the target molecule of the synthesis was compound **4** (Scheme 1). The material believed to have structure **4** was submitted by the synthetic chemist to the Spectroscopic Research Department for structural verification.

The sample preparation for NMR measurements was challenging because of the rather low solubility of chrysin derivatives in DMSO-d_6 and $CDCl_3$. In addition, the sample was probably contaminated with a small amount of insoluble inorganic salt; therefore, DMSO-d_6 and $CDCl_3$ were unable to dissolve the sample. An 80:20 (v/v) mixture of DMSO-d_6 and $CDCl_3$ dissolved most of the sample. The suspension was then filtered, resulting in a clear solution appropriate for NMR measurements.

 ©**AUTHOR(S)**



2 **4**

Scheme 1: The intended synthesis of compound **4**. Reagents and conditions: 1,4-dibromobutane, K_2CO_3, CH_3CN.

Initially we were in the consolidatory phase of the learning cycle regarding the structure elucidation of flavonoids. Therefore, we could quickly assign the characteristic signals in the NMR spectra to most of the atoms in the expected structure **4**. In the ^{1}H NMR spectrum of the sample, however, we observed a signal at 12.71 ppm which is characteristic of a hydrogen atom involved in a H-bond (Figure 3a, peak highlighted with red). This finding contradicts structure **4**. Thus, we assigned that signal to C(5)-OH based on our previous experience (Figure 3b). We suggested a monoalkylated analogue **5** as the product instead of the dialkylated compound **4**. According to the signals in the lower chemical shift range (0–5 ppm) of the ^{1}H NMR spectrum, we saw some evidence of alkylation taking place (Figure 3b), which was supported by the HMBC correlation between H_2-1' and C(7) (Figure 3c, green arrows).

This step in the structure elucidation was crucial, partly because we had come up with a new hypothesis for the molecular structure (creating a feeling of inventiveness and self-worth), and partly because it had sprung from our consolidated comfort zone of knowledge, providing a sense of structural confidence. Due to all of this, we could easily get obsessed with this conclusion. Having found a seemingly logical and appealing solution to a scientific problem (in this case, what is the structure of the main component?), it can become difficult to search further for any contradictory data that may potentially challenge the new idea. The active testing of one's new hypothesis is energy-consuming (mentally and often also instrumentally because new experimental data are required to exclude a faulty hypothesis). Also, it is seemingly not rewarding, because, most often, the original suggestion will prove to be right. Nevertheless, in principle, such extensive search is one of the essential traits of the scientific method, and is the hallmark of analytical thinking. Indeed, all of those cases – regardless of how few they are – that prove an apparently plausible and intellectually gratifying conclusion to be wrong justify such an extensive search! Moreover, finding a crack in our own argument via such careful and proactive pursuit, and, as a result, ultimately finding the *good* answer, can be a superbly uplifting moment, as discussed below.

First, we continued to look for further evidence supporting structure **5**. The COSY correlation between the ^{1}H resonances at 4.19 ppm and 2.04 ppm indicated the adjacency of H_2-1' and H_2-2' (Figure 3c, blue arrows). There was, however, no indication of any further groups on the chain according to the COSY spectrum. We could only observe a strong HMBC signal (^{1}H: 4.19 ppm, ^{13}C: 24.9 ppm) corresponding to a two-bond correlation of H_2-1' and C(2') (Figure 3c, green arrows). Still, we found two ^{1}H resonances [Figure 3a, 2.57 ppm and 1.25 ppm (highlighted with blue)] that were consistent with the remaining two CH_2 groups of **5** because the intensity of the signals agreed with the anticipated 2×2 hydrogen atoms. At this point, a less experienced interpreter (who may be prone to attribute the lack of those missing COSY cross-peaks to some measurement glitch to justify his/her novel idea) or any software that uses the ^{1}H NMR spectrum as the only input for structure validation, might claim that structure **5** is indeed the main component of the sample.

 ©AUTHOR(S)

We kept being suspicious, however, of the missing COSY and HMBC correlations involving the ^{1}H resonances at 2.57 ppm and 1.25 ppm. In addition, we found a peculiar HMBC correlation at ^{1}H: 2.04 ppm, ^{13}C: 24.9 ppm, suggesting a long-range correlation between H_2-2' and C(2'). Unless it was due to some measurement artefact, this cross-peak seemed to be inconsistent with the one-bond H_2-2'–C(2') connection in structure **5**, since such one-bond (direct) connections are normally detected in the HSQC spectrum, and should not give signals in the HMBC spectrum (Figure 3c, red arrows). We recognized that we were facing uncharted territory, i.e., entering the fourth phase of the learning cycle.

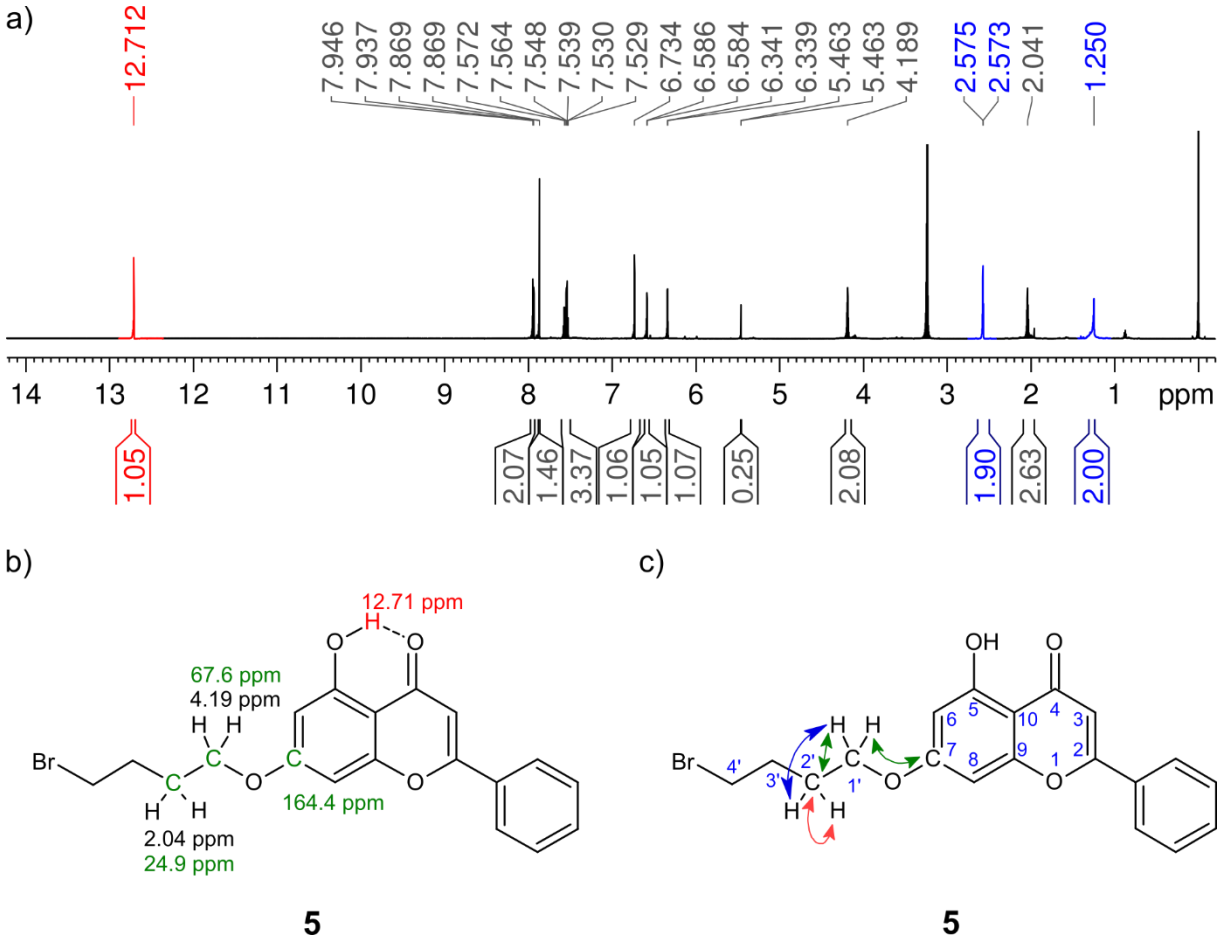


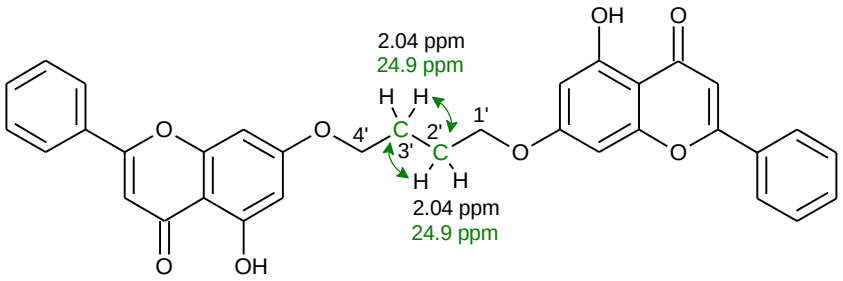
Figure 3: a) ^{1}H NMR spectrum of the sample (with highlighted resonances in red and blue). The chemical shift of the peaks is listed above the spectrum. Relative-signal intensities are shown below the abscissa; b) Our first proposed structure (**5**) instead of the expected structure (**4**). The ^{1}H NMR signal at 12.71 ppm belongs to the hydrogen atom shown in red; c) Observed correlations: COSY correlations are indicated with blue arrows, HMBC correlations with green and HSQC correlations with red arrows.

One possible way to resolve this contradiction is by closing the two-carbon-atom length chain with a heteroatom. By doing this, we must abandon the initial premise that the ^{1}H resonances at 2.57 ppm and 1.25 ppm belong to the main component. We can then interpret our observation as follows: The first signal belongs to under-deuterated solvent (CD_2H–S(=O)–CD_3) molecules (which usually appear at 2.50 ppm, however, the chloroform content of the solvent mixture shifted the signal by 0.07 ppm). The other signal (1.27 ppm) belongs to vacuum grease contamination. Its quantity in the investigated NMR sample may just have been unfortunate, giving rise to signals of deceptive intensity. The ^{1}H and ^{13}C NMR chemical shifts of C(2') were too low, however,

 ©**AUTHOR(S)**

(2.04 ppm and 24.9 ppm, respectively) to be connected with any possible heteroatom involved in the reaction (nitrogen, oxygen or a bromine atom). This was the point at which we were forced to make a mental leap, i.e., taking a completely new approach and revisiting previous decisions and interpretations. To perform the mental leap, we had to examine the HMBC spectrum more closely. We previously interpreted the HMBC cross peak at ^{1}H: 4.19 ppm, ^{13}C: 24.9 ppm as a two-bond correlation to support our structural proposal **5**. Could it actually be due, however, to a three-bond ^{1}H–^{13}C correlation? The HMBC experiment cannot distinguish between a two-bond and a three-bond correlation. By interpreting the aforementioned HMBC correlation as a three-bond H–C connection, we can continue the chain with another 24.9 ppm carbon atom (Figure 4). By assigning the 24.9 ppm chemical shift to two adjacent carbon atoms, a new idea came up, i.e., we may have a symmetrical molecule in our hands. The dihalide might have alkylated two molecules of chrysin, resulting in compound **6** (Figure 4).

As mentioned previously, we detected a peculiar HMBC cross-peak indicating a long-range (i.e., more than one bond) connection between the directly connected H_2-2' and C(2') atoms. Our newly proposed structure **6** could resolve that contradiction, i.e., the HMBC correlation was indeed a two-bond correlation, H_2-2'–C(3') (Figure 4). Later, the proposed structure **6** was confirmed by mass spectrometric measurements.



6

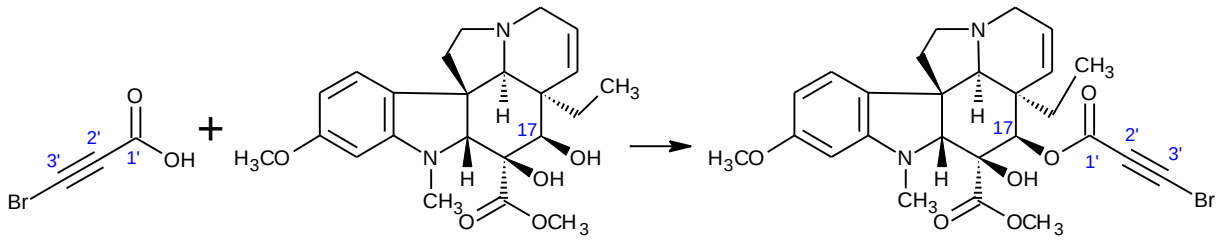
Figure 4: The correct structure (**6**) for the main component of the sample and interpretation of the ^{1}H: 2.04 ppm – ^{13}C: 24.9 ppm HMBC correlation.

As this example illustrates, cultivating the presented critical-thinking attitude and taking the appropriate mental leaps are the building blocks, and the driving force, of the spectroscopist's professional development.

Example 2

In the next example, we review the elucidation of a tricky and unexpected vindoline-related structure. The synthetic chemists attempted the esterification of 3-bromo-prop-2-yonic acid (Scheme 2, compound **7**) with 17-desacetylvindoline (**8**). The synthesis of reagent **7** was itself cumbersome (they obtained inseparable mixtures), so the synthetic chemists decided to use an impure, but apparently promising, batch (believing that **7** was its major component) for the synthesis of ester **9**. The product, after purification, was sent to us for structural investigation.

 ©**AUTHOR(S)**



7 **8** **9**

Scheme 2: Esterification of 3-bromo-prop-2-ynoic acid (**7**) with 17-desacetylvindoline (**8**) aiming for the synthesis of ester **9**. Reagents and conditions: *N*,*N*'-dicyclohexylcarbodiimide, 4-(*N*,*N*-dimethylamino)pyridine, dichloromethane.

The interpretation of the NMR spectra (^{1}H NMR, ^{13}C NMR, COSY, ROESY, HSQC, and HMBC) was, on the one hand, a relatively straightforward act within the mental realm of our already consolidated knowledge of showing that the vindoline skeleton had been unaffected during the synthesis. The HMBC experiment also indicated that acylation took place at the C(17)-O atom as the synthetic chemists had expected. The strong HMBC correlation at ^{1}H: 5.28 ppm, ^{13}C: 162.5 ppm corresponded to a three-bond correlation of H-17 and the carbon atom C(1') of the acyl group [Figure 5a, H-17–C(17)–O–C(1')]. The same HMBC correlation could be used to exclude the C(16)-O-acylated structure because four-bond HMBC correlations are generally weak or undetectable [Figure 5b, H-17–C(17)–C(16)–O–C(1')].

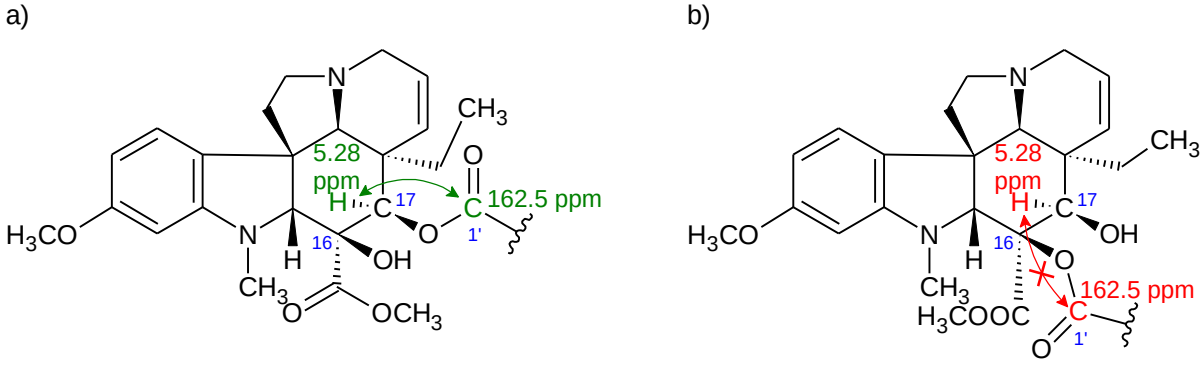


Figure 5: a) The HMBC correlation at ^{1}H: 5.28 ppm, ^{13}C: 162.5 ppm agrees well with the C(17)-O-acylated vindoline structure (a three-bond HMBC is expected); b) the same HMBC correlation contradicts the C(16)-O-acylated structure (a four-bond HMBC is often contradictory).

On the other hand, the structure elucidation of the prop-2-ynoyl group falls in the category of an "unfamiliar problem." Dealing with molecular fragments comprising very few hydrogen atoms can be troublesome, even for experienced spectroscopists, because most 2D NMR experiments rely on detecting the connectivity between a hydrogen atom (H) and another NMR-active isotope (X). There are, actually, surprisingly many structural misassignments in the literature[8] that can be traced back, for example, to the scarcity of available H–X connectivity data.[9]

For a novice spectroscopist, it is essential to be trained to avoid the mental trap of not making a properly conscientious distinction between a structure having been "certainly solved" and "seemingly solved". The latter can be all too easily confused with the former, especially when there is some deficiency in the available

 ©AUTHOR(S)

experimental spectral data. The extreme time pressures that are typical in industrial R&D & QC environments can much enhance the incidence of such confusion. Indeed, had it not been for our mental-trap-sensitive attitude, we could have erroneously validated the expected structure **9** since, subsequent to the assignment of all hydrogen and carbon atoms in the vindoline part of **9** to a corresponding ^{1}H NMR or ^{13}C NMR resonance, we noted that there were two additional ^{13}C NMR resonances (106.1 ppm and 128.1 ppm) that were consistent with the two remaining carbon atoms [C(2') and C(3'), Figure 5] in the structure.

Being mental-trap-conscious spectroscopists, we were not fully satisfied with our conclusions. For example, we suspected that for both C≡C carbon atoms the chemical shifts should be lower than 120 ppm. Therefore, we put some additional effort into spectrum interpretation and surprisingly found that one of the carbon atoms was connected to an additional H atom, as the HSQC correlation at ^{1}H: 7.21 ppm, ^{13}C: 128.1 ppm indicated. However, the singlet in the ^{1}H NMR spectrum at 7.21 ppm showed no correlations to any other H or C atoms in the HMBC, COSY, and ROESY spectra. Furthermore, we could not detect any connection of the ^{13}C resonance at 106.1 ppm to any hydrogen atom. We seemed to have an, apparently isolated, CH group in our hands, and an isolated quaternary carbon atom. Some pieces of the puzzle were certainly missing.

We could argue that a proton resonance with no correlations in the COSY and the HMBC spectrum could be due to an impurity, even when it has an almost integer-relative intensity (the intensity of the singlet at 7.21 ppm is 0.95 when a hydrogen atom in vindoline is calibrated to an intensity of 1.00), as presented in Example 1 (cf. the peaks due to grease and under-deuterated solvent molecules). However, we were not able to come up with a plausible suggestion for the structure of the impurity with such chemical shifts (^{1}H: 7.21 ppm, ^{13}C: 128.1 ppm and 106.1 ppm).

We desperately needed a mental leap. To that end, we had to go against a quite new, yet unconsolidated piece of knowledge we had acquired during the solving of Example 1. We had to consider *not* attributing the missing 2D NMR correlations as being due to an impure material. As an alternative approach, we assumed that we were facing a very curious situation: the coupling constants between some resonances of the molecule were almost zero, which prohibited the detection of their connection. In pursuit of that hypothesis, we decided to spend additional instrumental effort aimed at detecting the suspected very small ^{1}H–^{13}C couplings. We reran the HMBC experiment with some adjustments: the HMBC is conventionally optimized to the detection of $J_{C,H}$ ≈ 8 Hz; we modified that optimum to $J_{C,H}$ ≈ 2 Hz. This is a well-known technique for the detection of long-range (usually 4 or 5-bond) HMBC correlations, but it comes at the cost of a worse signal-to-noise ratio (S/N). As suspected, the ^{1}H resonance of the missing CH group at 7.21 ppm gave HMBC correlations to carbon atoms at 106.1 ppm (the other, at first sight, isolated carbon atom) and 162.5 ppm (the acyl group connected to C(17)-O atom of vindoline). We found three possible structures (Figure 6, **10**–**12**) that were consistent with the 2 Hz HMBC data. Taking into account the chemical context, we suspected that X was probably a Br atom.

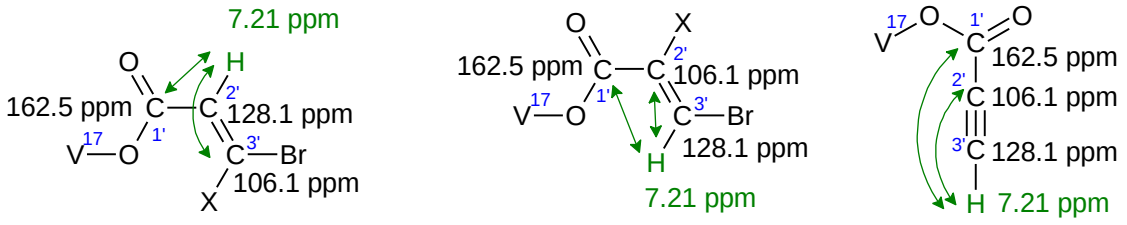


Figure 6. Three structures (**10**–**12**) that were consistent with the 2 Hz HMBC data (green arrows correspond to the detected multiple-bond H–C connections); V: vindoline unit, excluding the C(17)-O atom; X: any group without H or C atoms.

 ©**AUTHOR(S)**

Structure **12** can be excluded as follows: the ^{1}H NMR chemical shift of an alkyne H atom is highly unlikely to exceed 5 ppm; the 7.21 ppm shift does not correspond to a prop-2-ynoyl group. To decide between **10** and **11**, we used the 1,1-ADEQUATE (Adequate Sensitivity Double-Quantum Spectroscopy) experiment. The 1,1-ADEQUATE experiment detects spin systems consisting of directly connected ^{1}H–^{13}C–^{13}C atoms. In the resulting spectrum, we can correlate ^{1}H and ^{13}C resonances when there is one ^{13}C atom between them. Therefore, the ambiguity (two-bond H–C correlations cannot be differentiated from three-bond H–C correlations) associated with the HMBC experiment can be resolved. Due to the low probability (about 10^{-4}) of finding two adjacent ^{13}C nuclei, however, the sensitivity of the ADEQUATE experiment is rather low. It took 10 h of measurement time from a sample of ca. 13 mmol/L on an 800 MHz spectrometer with a cryoprobe.

In the 1,1-ADEQUATE spectrum of the sample (Figure 7a), we detected a peak at ^{1}H: 7.21 ppm (H-2'), ^{13}C: 162.5 ppm [C(1')], meaning that there is one ^{13}C atom [the C(2')] between H-2' and C(1') (Figure 7b). The new piece of information corresponds with structure **10**, but not with **11**. According to subsequent high-resolution mass spectroscopic measurements (M+H = 625.05487 Da, $C_{26}H_{31}O_6N_2Br_2$), we concluded that X was a bromine atom [C(3')-Br_2].

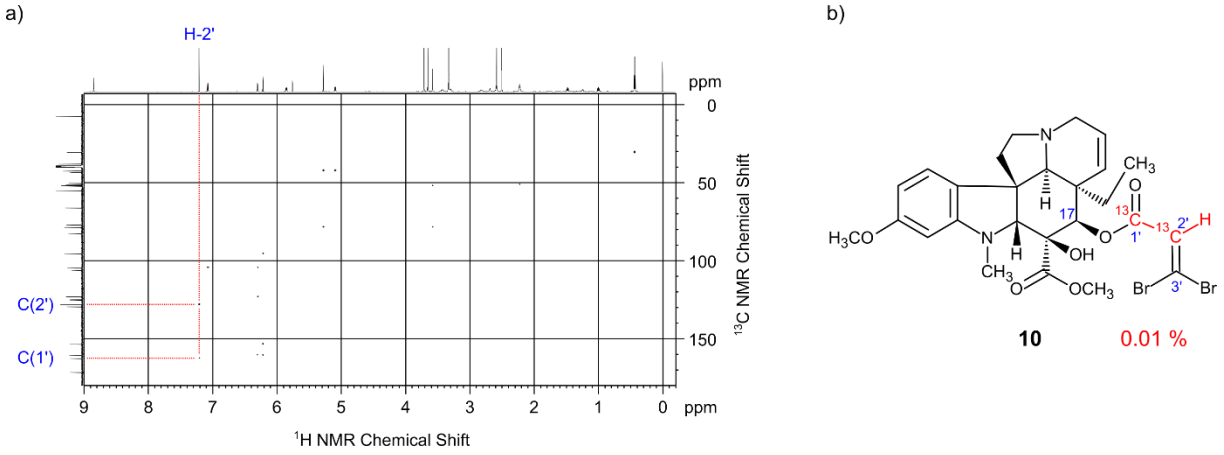


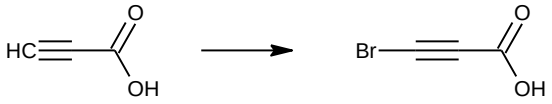
Figure 7. a) The 1,1-ADEQUATE spectrum of **10**; the peak at ^{1}H: 7.21 ppm, ^{13}C: 162.5 ppm [H-2'–C(1')] was barely detectable (S/N = 4); b) interpretation of the H-2'–C(1') ADEQUATE peak; The peak at ^{1}H: 7.21 ppm, ^{13}C: 128.1 ppm [H-2'–C(2')] is an HSQC artifact.

Example 3

From the perspective of a synthetic chemist, considering the reactants **7** and **8** (Scheme 2), the elucidated structure **10** in Example 2 is quite questionable. There is no apparent rationale behind the formation of the dibromo compound **10** by the esterification of 3-bromoprop-2-ynoic acid **7** with an alcohol (17-desacetylvindoline, **8**).

To resolve the discrepancy, we present the challenges of the analysis of a sample related to the preparation of reagent **7**. 3-Bromoprop-2-ynoic acid was prepared via the addition of a bromine radical to prop-2-ynoic acid **13** (Scheme 3).

 ©**AUTHOR(S)**



13 **7**

Scheme 3: Intended preparation of **7** from **13**. Reagents and conditions: *N*-bromosuccinimide, $AgNO_3$, acetone.

The first challenge in the structure elucidation relates to contamination, which makes the identification of the expected compound **7** very difficult. The sample was also prone to degradation in DMSO-d_6. In the ^{1}H NMR spectrum (Figure 8a) of a newly prepared solution of the sample, we readily identified large amounts of succinimide (**14**, NH: 11.05 ppm, CH_2: 2.56 ppm; the relative intensity is calibrated to 1.00 per H atom). We could identify the corresponding ^{13}C NMR resonances (CH_2: 29.4 ppm, C=O: 179.3 ppm, Figure 8b and 8c) via the ^{1}H–^{13}C HSQC and the ^{1}H–^{13}C HMBC experiment.

There were six additional peaks in the ^{13}C NMR spectrum of the sample (Figure 8b, 54.1, 73.4, 103.9, 129.7, 152.9, 164.0 ppm) that we could not interpret yet (the peak at 0.0 ppm belongs to TMS; the peak at 39.4 ppm belongs to DMSO-d_6).

Unfortunately, the expected compound **7** does not give any sharp signals in the ^{1}H NMR spectrum. The only expected signal is a broad peak around 13 ppm (COOH), which often melts into the water peak (a characteristic feature of an exchangeable, acidic proton). Consequently, **7** gives no signals in the ^{1}H–^{13}C HSQC and the ^{1}H–^{13}C HMBC spectrum which otherwise could have been used for collecting the ^{1}H NMR and the ^{13}C NMR resonances that belong to the same molecule.

We found another impurity in a molar ratio of 0.65 (based on ^{1}H NMR relative intensities, Figure 8a), giving rise to the singlet resonance at ^{1}H: 7.20 ppm. Based on HSQC and HMBC data, we found that two peaks (^{13}C: 129.7 ppm and 103.9 ppm) belonged to the impurity. After having found the correct structure in Example 2, we became familiar with those chemical shifts (cf. **10**, Figure 6) and applied (Figure 1, "Applying") the new piece of knowledge for proposing structure **15**. Based on the analogy with **10**, we surmised that the ^{13}C NMR resonance at 164.0 ppm belonged to the carboxylic acid group of **15**. Despite the fact that the 164.0 ppm resonance gave no correlations in the HMBC spectrum, we were content with our structural proposal because we have learned in Example 2 that two-bond ^{1}H–^{13}C couplings can be very small for a 3,3-dibromoprop-2-enoyl moiety.

There were three ^{13}C resonances (54.1, 73.4, and 152.9 ppm) remaining. To find out whether **7** is one of the main components of the sample, we had to learn a technique with which we were yet unfamiliar (Figure 1, "Learning"). We started to introduce molecular modeling and quantum-mechanical (QM) NMR chemical-shift calculations into our toolkit at a basic level. We performed calculations for each suspected components of the sample (**14**, **15**, and **7**); the QM-predicted ^{13}C NMR shifts are presented in Figure 8d.

We note that the interpretation of the QM results has its own associated pitfalls; to avoid them, thorough training is essential. For example, the chemical shift of the carbon atoms connected to bromine atoms is systematically overestimated (by of the order of 30 ppm per C–Br bond, according to our observations), as indicated by "**" in Figure 8d. Considering these systematic deviations in QM chemical shift calculations, the calculated and observed chemical shifts are comparable overall (less than 4 ppm disagreement for sp^3 and sp^2 carbon atoms; less than 8 ppm difference for C(2) in **7**). In conclusion, QM calculations support our structural proposals. Compounds **15** and **7** were later detected in MS measurements (nominal mass M=148; C_3HO_2Br, and nominal mass M=228; $C_3H_2O_2Br_2$).

 ©**AUTHOR(S)**

The final challenge was the determination of the amount of the expected compound **7** in the sample. As **7** does not give any sharp ^{1}H NMR resonances required for reliable peak integration (from which the molar ratio of the components can be calculated), we could only rely on ^{13}C NMR peak intensities. Due to the low S/N and relatively fast pulsing in ^{13}C NMR (as compared to the T_1 relaxation time of ^{13}C resonances), peak intensities correlate with the molar ratio poorly under routine measurement conditions. We could only estimate the ratio of **15** and **7**, which was about 2:1, based on a comparison of the intensity of the COOH peaks (152.9 ppm and 164.0 ppm). A more accurate estimation (via quantitative ^{13}C NMR measurements) would have required larger amounts of sample and long experiment times (several hundreds of scans using a relaxation delay on the order of half a minute). The experiment was not feasible due to the low quantity of the sample (several thousands of scans, corresponding to at least a day of measurement, would have been needed to achieve a quantitative ^{13}C NMR spectrum with adequate S/N).

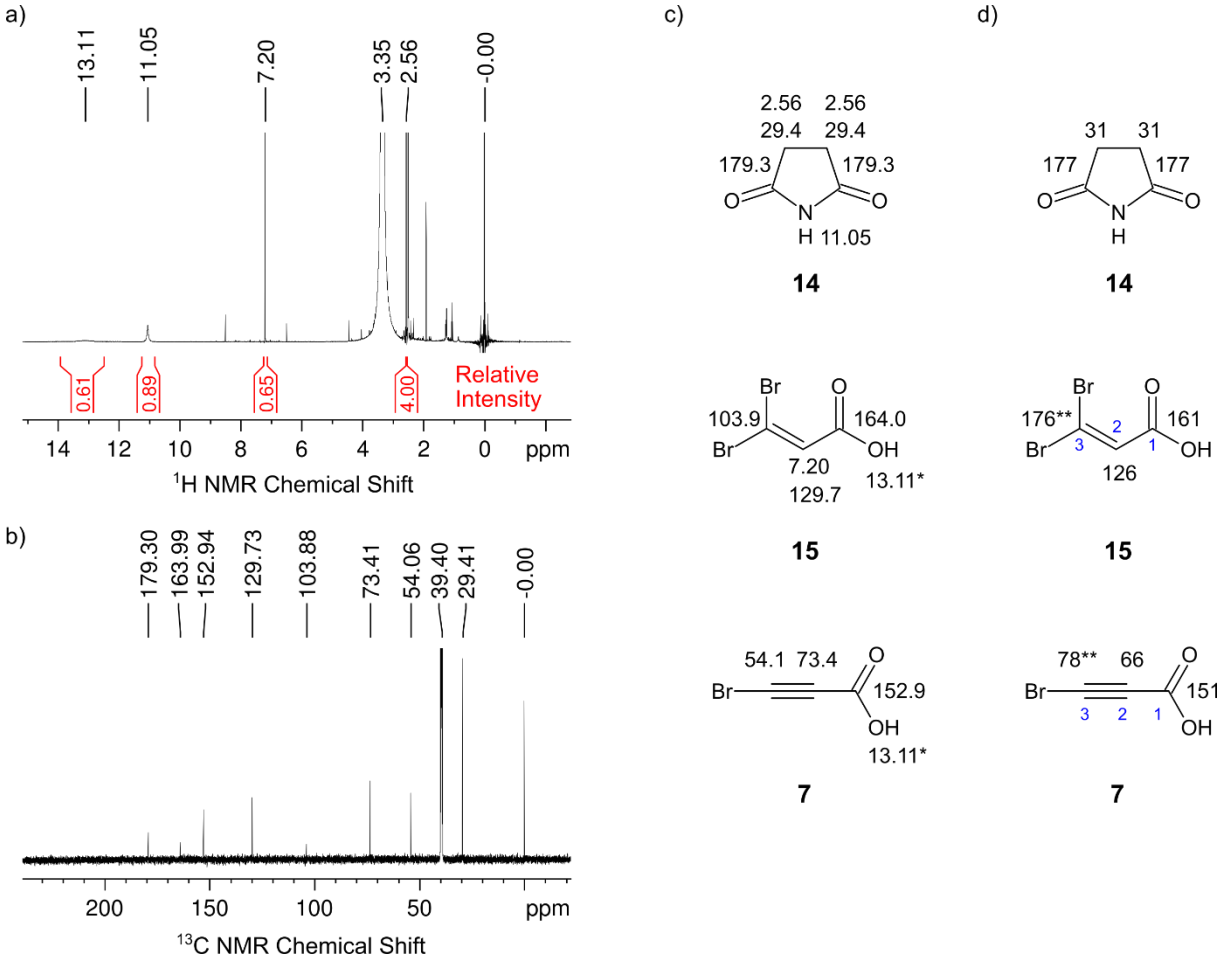


Figure 8. a) ^{1}H NMR spectrum of the sample; b) ^{13}C NMR spectrum of the sample; c) Resonance assignment of the main components **14**, **15,** and **7** (*: broad peak, uncertain chemical shift); d) Calculated ^{13}C NMR chemical shifts at the B3LYP-D3/6-31+G(d,p) level of theory (**: unreliable due to C–Br bonds).

 ©**AUTHOR(S)**

Conclusions

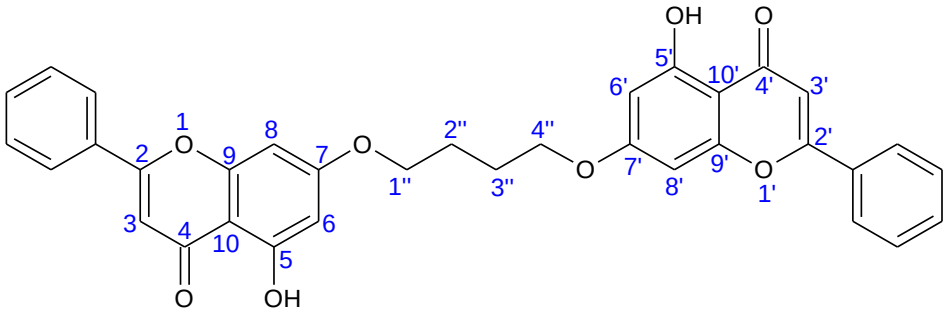
We have presented our notion of how scientific expertise develops via an evolutionary process involving learning, applying, consolidating, stumbling across unfamiliar problems, running into mental traps, and executing mental leaps. We have demonstrated this course via three intriguing examples of correct structure elucidation using NMR spectroscopy. We pointed out that there is plenty of room for pitfalls associated with the interpretation of NMR spectra. Critical thinking, mental-trap-consciousness, as well as being brave enough to challenge previously consolidated knowledge, is essential to avoid coming to faulty structural conclusions. Striving for high-confidence structure elucidation can be demanding, both mentally and instrumentally, especially under high time-pressure conditions, however, it is rewarding in the long run. In our opinion, the discussed learning model and scientific *modus operandi* are relevant to other practical, or even theoretical fields as well.

Experimental Section

General. ^{1}H And ^{13}C NMR measurements were performed on a Bruker Avance III HDX 500 MHz NMR spectrometer equipped with a ^{1}H {^{13}C/^{15}N} 5 mm TCI CryoProbe, or a Bruker Avance III HDX 800 MHz NMR spectrometer equipped with a ^{1}H-^{19}F{^{13}C/^{15}N} 5 mm TCI CryoProbe (Bruker Corporation, Billerica, MA, USA). ^{1}H and ^{13}C chemical shifts are given on the delta scale as parts per million (ppm) with tetramethylsilane (TMS) as reference. d_6-Dimethyl sulfoxide (DMSO-d_6), *d*-chloroform ($CDCl_3$) or their mixture was used as solvent. ^{1}H–^{1}H scalar spin-spin connectivity, direct ^{1}H–^{13}C connectivity, and long-range ^{1}H–^{13}C connectivity were established from two-dimensional (2D) COSY, HSQC and HMBC experiments, respectively. ^{1}H–^{1}H spatial proximities were determined by using the 2D ROESY experiment. NMR spectra were processed with Bruker TopSpin 3.5 pl 6 (Bruker Corporation, Billerica, MA, USA) and ACD/Spectrus Processor version 2017.1.3 (Advanced Chemistry Development, Inc., Toronto, ON, Canada).

^{1}H And ^{13}C NMR characterization of the compounds

Compound ***6***



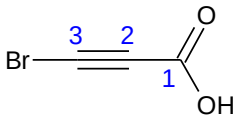
^{1}H NMR (799.7 MHz; DMSO-d_6 : $CDCl_3$ 80:20 (v/v)): δ (ppm) 2.04 (4H; br m; H_2-2", H_2-3"); 4.19 (4H; br t; $^{3}J_{HH}$ 4.9 Hz; H_2-1", H_2-4"); 6.34 (2H; d; $^{4}J_{HH}$ 1.7 Hz; H-6, H-6'); 6.59 (2H; d; $^{4}J_{HH}$ 1.7 Hz; H-8, H-8'); 6.73 (2H; s; H-3, H-3'); 7.52-7.59 (6H; m; C(2)-Ph: H_{meta}, C(2)-Ph: H_{para}, C(2')-Ph: H_{meta}, C(2')-Ph: H_{para}); 7.93-7.96 (4H; m; C(2)-Ph: H_{ortho}, C(2')-Ph: H_{ortho}); 12.71 (2H; s; C(5)-O<u>H</u>, C(5')-O<u>H</u>).

^{13}C NMR (201.1 MHz; DMSO-d_6): δ (ppm) 24.9 (C-2", C-3"); 67.6 (C-1"; C-4"); 92.6 (C-8, C-8'); 98.1 (C-6, C-6'); 104.9 (C-10; C-10'); 105.1 (C-3; C-3'); 125.8 (C(2)-Ph: C_{ortho}, C(2')-Ph: C_{ortho}); 128.6 (C(2)-Ph: C_{meta}, C(2')-Ph: C_{meta});

 ©**AUTHOR(S)**

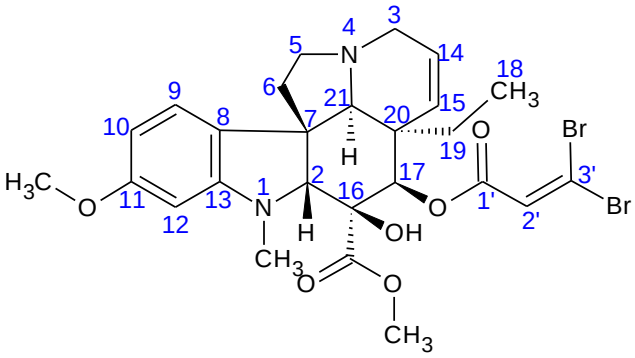
130.5 (C(2)-Ph: C_{ipso}, C(2')-Ph: C_{ipso}); 131.4 (C(2)-Ph: C_{para}, C(2')-Ph: C_{para}); 157.2 (C-9, C-9'); 161.3 (C-5, C-5'); 163.3 (C-2, C-2'); 164.4 (C-7, C-7'); 181.7 (C-4, C-4').

Compound ***7***



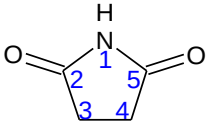
^{13}C NMR (125.7 MHz; DMSO-d_6): δ (ppm) 54.1 (C-3); 73.4 (C-2); 152.9 (C-1).

Compound ***10***



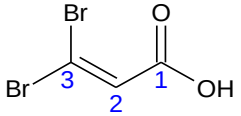
^{1}H NMR (499.9 MHz; DMSO-d_6): δ (ppm) 0.43 (3H; t; $^{3}J_{HH}$ 7.3 Hz; H_3-18); 0.99 (1H; dq; $^{2}J_{HH}$ 14.0 Hz; $^{3}J_{HH}$ 7.3 Hz; H_x-19); 1.47 (1H; dq; $^{2}J_{HH}$ 14.0; $^{3}J_{HH}$ 7.3 Hz; H_y-19); 2.18-2.25 (2H; m; H_2-6); 2.56-2.60 (4H; N(1)-C<u>H</u>$_3$, H_x-5); 2.68 (1H; s; H-21); 2.77-2.85 (1H; m; H_x-3); 3.26-3.31 (1H; m; H_y-5); 3.42 (1H; br dd; $^{2}J_{HH}$ 16.0 Hz; $^{3}J_{HH}$ 4.8 Hz; H_y-3); 3.57 (1H; s; H-2); 3.65 (3H; s; C(16)-COOC<u>H</u>$_3$); 3.71 (3H; s; C(11)-OC<u>H</u>$_3$); 5.06-5.11 (1H; m; H-15); 5.28 (1H; s; H-17); 5.85 (1H; ddd; $^{3}J_{HH}$ 10.2 Hz; $^{3}J_{HH}$ 4.9 Hz; $^{3}J_{HH}$ 1.4 Hz; H-14); 6.22 (1H; d; $^{4}J_{HH}$ 2.2 Hz; H-12); 6.30 (1H; dd; $^{3}J_{HH}$ 8.2 Hz; $^{4}J_{HH}$ 2.2 Hz; H-10); 7.07 (1H; d; $^{3}J_{HH}$ 8.2 Hz; H-9); 7.21 (1H; s; H-2'); 8.84 (1H; s; C(16)-O<u>H</u>).
^{13}C NMR (125.7 MHz; DMSO-d_6): δ (ppm) 7.5 (C-18); 30.5 (C-19); 38.0 (N(1)-<u>C</u>H$_3$); 42.4 (C-20); 43.5 (C-6); 50.3 (C-3); 51.1 (C-5); 51.8 (C(16)-COO<u>C</u>H$_3$); 52.0 (C-7); 55.0 (C(11)-O<u>C</u>H$_3$); 66.1 (C-21); 76.9 (C-17); 78.6 (C-16); 82.7 (C-2); 95.5 (C-12); 104.6 (C-10); 106.1 (C-3'); 123.1 (C-9); 124.8 (C-14); 125.2 (C-8); 128.1 (C-2'); 129.4 (C-15); 153.4 (C-13); 160.5 (C-11); 162.5 (C-1'); 171.5 (C(16)-<u>C</u>OOCH$_3$).

Compound ***14***



^{1}H NMR (499.9 MHz; DMSO-d_6): δ (ppm) 2.56 (4H; s; H_2-3, H_2-4); 11.05 (1H; br; N(1)-<u>H</u>).
^{13}C NMR (125.7 MHz; DMSO-d_6): δ (ppm) 29.4 (C-3, C-4); 179.3 (C-2, C-5).

Compound ***15***



 ©**AUTHOR(S)**

^{1}H NMR (499.9 MHz; DMSO-d_6): δ (ppm) 7.20 (1H; s; H-2).
^{13}C NMR (125.7 MHz; DMSO-d_6): δ (ppm) 103.9 (C-3); 129.7 (C-2); 164.0 (C-1).

Acknowledgements

The authors wish to thank Viktória Donkó-Tóth, Szabolcs Mayer and László Hazai (Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics) for some of the chemical synthesis, and Miklós Dékány (Spectroscopic Research Department, Gedeon Richter Plc.) for the HRMS measurements.

Supplementary Material

The ^{1}H and the ^{13}C NMR, as well as the 2D NMR spectra of the investigated samples are provided in the Supplementary Material.

References

1. Szántay, C. Jr. In *Anthropic Awareness: The Human Aspects of Scientific Thinking in NMR Spectroscopy and Mass Spectrometry*; Szántay, C. Jr. Ed.; Elsevier: New York, 2015; pp 3-93.
https://doi.org/10.1016/B978-0-12-419963-7.00001-8
2. Szántay, C. Jr. In *Anthropic Awareness: The Human Aspects of Scientific Thinking in NMR Spectroscopy and Mass Spectrometry*; Szántay, C. Jr. Ed.; Elsevier: New York, 2015; pp 231-256.
https://doi.org/10.1016/B978-0-12-419963-7.00006-7
3. Szántay, C. Jr.; Moser, E. *Front. Phys.* **2018**, *6*, 68.
https://doi.org/10.3389/fphy.2018.00068
4. Szántay, C. Jr.; Moser, E. *Front. Phys.* **2018**, *6*, 70.
https://doi.org/10.3389/fphy.2018.00068
5. Elyashberg, M.; Williams, A. J.; Blinov, K. *Nat. Prod. Rep.* **2010**, *27,* 1296–1328.
https://doi.org/10.1039/c002332a
6. Dhyani, P.; Quispe, C.; Sharma, E. *et al. Cancer Cell Int* **22**, 206 (2022).
https://doi.org/10.1186/s12935-022-02624-9
7. Mayer, S.; Nagy, N.; Keglevich, P.; Szigetvári, Á.; Dékány, M.; Szántay, C. Jr.; Hazai, L. *Chem. Biodiversity* **2022**, *19*, e202100725.
https://doi.org/10.1002/cbdv.202100725
8. Suyama, T. L.; Gerwick, W. H.; McPhail, K. L. *Bioorg. Med. Chem.* **2011**, *19*, 6675-6701.
https://doi.org/10.1016/j.bmc.2011.06.011
9. Martin, G. E.; Buevich, A. V.; Reibarkh, M.; Singh, S. B.; Ondeyka, J. G.; Williamson R. T. *Magn. Reson. Chem.* **2013**, *51*, 383-390.
https://doi.org/10.1002/mrc.3952

 ©**AUTHOR(S)**

This paper is an open access article distributed under the terms of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/)