

Nuclear Magnetic Resonance in the Structural Elucidation of Natural Products

William F. Reynolds and Eugene P. Mazzola

Contents

1	Introduction.....	224
2	Dereplication: Distinguishing Between New and Known Natural Products.....	225
3	Quantitative NMR.....	228
4	Using 2D NMR to Determine Skeletal Structures of Natural Products.....	230
5	Avoiding Getting the Wrong Skeletal Structure.....	233
6	Determination of Configuration and/or Conformation of Natural Products.....	235
7	An Example of a Solved Structure: Kauradienoic Acid.....	238
7.1	HSQC Data.....	239
7.2	COSY and TOCSY Data.....	240
7.3	HMBC Data.....	244
7.4	General Molecular Assembly Strategy.....	246
7.5	A Specific Molecular Assembly Procedure.....	248
7.6	Determination of Overall Stereochemistry and Proton Chemical Shift Assignments.....	252
8	Computer-Assisted Structure Elucidation.....	257
8.1	Guyanin.....	258
8.2	T-2 Toxin.....	262
8.3	Kauradienoic Acid.....	267
9	The Effect of Dynamic Processes on the Appearance of NMR Spectra of Natural Products and Other Organic Compounds.....	268
10	The Relative Advantages and Disadvantages of Different Pulse Sequences.....	273
11	Liquid-Chromatography–NMR.....	277

W.F. Reynolds (✉)

Department of Chemistry, University of Toronto, Toronto, ON, Canada M5S 3H6

e-mail: wreynold@chem.utoronto.ca

E.P. Mazzola

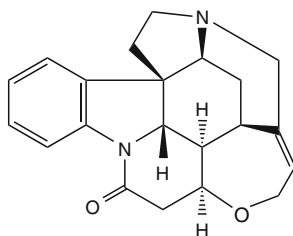
University of Maryland-FDA Joint Institute, College Park, MD, 20742 USA

e-mail: emazzola@umd.edu

12	Probe Choices	278
12.1	Essential Probe Features for Natural Product Research.....	278
12.2	Ambient-Temperature Probes.....	279
12.3	Cryogenically Cooled Probes.....	279
12.4	Microprobes.....	280
13	A Fully Automated Setup of 2D NMR Experiments for Organic Structure Determination.....	281
14	Parameter Choices for Acquisition and Processing of 1D and 2D NMR Spectra.....	283
14.1	Basics of NMR Data Acquisition.....	283
14.2	Recommended Acquisition and Processing Parameters for 1D Spectra.....	286
14.3	Basics of 2D NMR.....	290
14.4	Recommended Acquisition and Processing Parameters for Commonly Used 2D Experiments and Selective 1D Experiments.....	298
15	Conclusions.....	303
	References.....	303

1 Introduction

For at least the first 60 years of the twentieth century, natural product structure elucidation was a lengthy and extremely difficult task, usually involving numerous synthetic and degradative reactions. As such, it attracted the attention of many of the world's leading synthetic organic chemists. However, since then, dramatic improvements in nuclear magnetic resonance (NMR) spectroscopy, along with similar advances in mass spectrometry, X-ray crystallography, and chromatography, have totally revolutionized natural product structure elucidation. For example, although strychnine (**1**) had been purified from its plant source over a century earlier, it took more than 40 years of effort by several research groups, before its structure was finally determined by Sir *Robert Robinson* in 1946 (1) and later confirmed by a total synthesis of strychnine by *R.B. Woodward* and co-workers (2). Now with modern NMR methods and a state-of-the-art NMR spectrometer, the total structure of strychnine could be determined in 24 h using less than 1 mg of sample.



1

This is partly due to dramatic improvements in spectrometers. For example, one of the authors (*WFR*) got his start in NMR in 1960, using a 60-MHz spectrometer with a guaranteed proton signal/noise (*S/N*) of 10:1 for 1.0% ethylbenzene. Now spectrometers are available up to 1,000 MHz, and *S/N* specifications on

cryogenically cooled probes are approaching 10,000:1 for 0.1% ethylbenzene! However, developments in methodology have been equally critical. Of these, the development of *Fourier* transform (FT) NMR by *Wes Anderson* and *Richard Ernst* in 1966 (3) and two-dimensional (2D) NMR by *Richard Ernst* in 1976 (4) were particularly important, with *Ernst* receiving the 1991 *Nobel Prize* in Chemistry for his contributions (5).

This chapter is written with the goal of helping natural product chemists use modern NMR methods as effectively as possible for natural product structure elucidation. It is assumed that the reader will have at least a basic knowledge of the use of NMR in organic chemistry, at the level covered in senior undergraduate or graduate courses in spectroscopic methods for organic structure determination and as provided by texts such as *Silverstein et al.* (6) or *Lambert et al.* (7). Therefore, topics such as typical values of chemical shifts and coupling constants and factors affecting these parameters will not be discussed. Instead, emphasis will be placed on the rapid determination as to whether an isolated compound is known or new, the information content of different 2D and selective 1D experiments, their use in combination for structure elucidation, possible pitfalls in structure determination by NMR and how these can be avoided or overcome, and the use of computer-assisted structure elucidation (CASE). Considerable space will also be devoted as to how to make the correct choices of acquisition parameters and data processing methods and parameters. While this topic has been discussed in two books (8, 9) and at least two review articles (10, 11), the importance of this topic seems to be underappreciated by most users of NMR. In this regard, it is important to recognize that the default parameters provided in the manufacturers' software packages or a widely used book, which provides default parameters for many NMR experiments (8), may not be ideal and that sometimes dramatic improvements in results can be obtained by different choices (12).

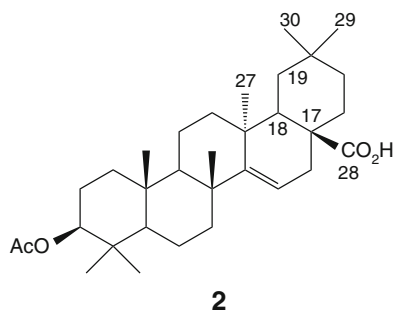
The references given in the various sections are intended to be representative rather than comprehensive and are often chosen from our own published work. Finally, in places the authors refer to specific NMR spectrometer manufacturers and their products. This is done so that the reader is aware of various options available and the different procedures that sometimes have to be followed in processing or interpreting the data obtained on different spectrometers due to hardware and/or software differences. It should not be taken as indicating a preference of one spectrometer over another.

2 Dereplication: Distinguishing Between New and Known Natural Products

Much of current natural product research involves bioassay screening of crude chromatographic fractions, followed by separation of fractions with promising activity into pure compounds for further detailed testing. Both in terms of time and cost, it is important to be able to quickly identify any known compounds to avoid wasting valuable analytical instrument time on detailed characterization of known

compounds. However, it is also important that this identification process be sufficiently reliable so that there is minimal risk of mistaking an unknown compound for a known compound. Unfortunately, these two requirements conflict, forcing one to compromise either speed or reliability to a certain extent. The advantages and disadvantages of the various approaches that can be used are discussed below.

In view of its intrinsic high sensitivity, mass spectrometry (MS) is a logical initial choice for dereplication, most commonly in conjunction with liquid chromatography (see this volume, *Budzikiewicz H (2014) Mass Spectrometry in Natural Product Structure Elucidation. Progr Chem Org Nat Prod 100:77*). Only a small fraction of the crude extract will often be sufficient for an LC-MS investigation, using either low (LR) or high (HR) resolution MS. While LR-MS is quicker, it is not sufficiently accurate to determine exact molecular formulae, and there are often a large number of structures consistent with the parent ion. HR-MS will give the empirical formulae for a much smaller number of possible structures. Even in the ideal situation where only one empirical formula is consistent with the exact mass within acceptable error limits (usually 5 ppm), there can still be a number of isomeric structures consistent with this mass. However, MS databases (see this volume, *Budzikiewicz H (2014)* will usually list all known structures consistent with an empirical formula. In principle, one may be able to use MS fragmentation patterns to favor one structure over the others. However, the fragmentation pattern and/or the relative intensities of fragment ions can be different, depending on the ionization source and/or ionization energy. In addition, there can be ambiguities in the interpretation of fragmentation patterns. For example, a number of years ago, we determined the structure of a triterpene, 3-acetylaleuritic acid (**2**), by 2D NMR and discovered that three different structures (one of which was correct) had been proposed for the same compound, mainly based on different interpretations of the MS fragmentation pattern (*13*).



An alternative, intermediate sensitivity, approach is to use ^1H NMR spectroscopy for dereplication. While in principle this can be done by LC-NMR in flow mode, it is better to collect the samples in solid phase extraction cartridges and then transfer them to NMR tubes (see Sect. 11). Ideally, spectra should be obtained using either a cryogenically cooled probe or microprobe for maximum sensitivity (see Sect. 12),

but unfortunately these probes are not available to many natural product groups. However, even using 5-mm tubes in an ambient temperature probe, one can get adequate ^1H spectra in no more than one minute with 1 mg of sample. The main problem with ^1H NMR for dereplication purposes is that the appearance of these spectra can change significantly with spectrometer frequency and solvent. Thus, it may be difficult to determine with certainty whether a particular spectrum is identical to one from a database, which was obtained under different experimental conditions. However, several open access and commercial databases also include programs, that will predict the appearance of a ^1H spectrum for a given structure at a given frequency (NMRWiki provides a list of NMR databases (14)). Thus, if one has a list of candidate structures from HR-MS, one can compare the calculated spectra for these compounds with an experimental spectrum, hopefully leading to the most probable structure.

Clearly, the most reliable identification of a known compound would be provided by a combination of a ^{13}C NMR spectrum and a HR-MS spectrum. For that reason, journals increasingly require both a good quality ^{13}C spectrum and a HR-MS spectrum when reporting a new organic compound. Unfortunately, however, the intrinsic low sensitivity of ^{13}C NMR often makes it impractical to use a full ^{13}C spectrum for dereplication purposes unless one is fortunate enough to have access to a cryogenically cooled probe optimized for ^{13}C detection. However, a good compromise is to obtain a DEPT-135 spectrum (Sect. 14) or an edited HSQC spectrum (Sect. 10). Both spectra give ^{13}C data for all protonated compounds with CH and CH_3 peaks of opposite phase to CH_2 peaks and can be obtained in similar time and significantly more quickly than for a full ^{13}C spectrum (10). However, the HSQC spectrum has the additional major advantage of providing the chemical shifts of the proton(s) attached to each specific carbon. In addition, the use of non-uniform sampling along the evolution axis (15) has the potential to further increase the sensitivity of HSQC spectra (16).

Even greater time saving can be achieved by using either the SOFAST-HMQC sequence of *Brutscher* (17) or the ASAP-HMQC sequence of *Kupce* and *Freeman* (18). Both of these sequences allow the use of far shorter than usual relaxation delays, dramatically reducing the time to acquire a 2D ^1H - ^{13}C correlation spectrum. For example, it has recently been shown that one can obtain a high quality HMQC spectrum on a 400-MHz spectrometer in under one minute for a 5 mg sample of a compound of over 400 molecular weight, using the ASAP-HMQC sequence (19). The only disadvantages of these sequences are that they do not provide spectral editing and have poorer ^{13}C resolution than provided by HSQC spectra. Nevertheless, they have tremendous potential, particularly for rapid screening and dereplication.

It is unfortunate that, so far as we are aware, none of the existing $^1\text{H}/^{13}\text{C}$ databases are integrated to take advantage of the additional information provided by HSQC or HMQC, *i.e.* they do not correlate ^1H chemical shifts with the ^{13}C chemical shifts of the carbons to which the protons are bonded. However, we understand that this integration is in progress and, when integrated $^1\text{H}/^{13}\text{C}$ databases are available, this in conjunction with HR-MS will provide a highly reliable approach for dereplication.

One problem with the use of NMR databases for dereplication is that full spectral assignments are not available for many older known compounds. Unfortunately, natural product journals usually will not allow publication of assignments for known compounds unless one has data for a significant number of related compounds. However, several open access databases will accept these data (14). Thus, we would strongly encourage natural product chemists to carry out full assignments for known compounds, where these are not available, and deposit the information in at least one of these databases, for the benefit of all in the field. Procedures for making these assignments, using combinations of 2D NMR spectra, are discussed in Sect. 4. The same procedures can be used to identify the structures and fully assign the spectra of new compounds.

3 Quantitative NMR

One very important advantage of ^1H NMR over all of the other types of spectroscopy used in natural product chemistry is that it is intrinsically quantitative. To realize this advantage, it is necessary to take reasonable care in the choice of acquisition parameters, but this is not difficult (see Sect. 14 for a discussion of parameter choices). There has been a recent significant increase in interest in the use of quantitative NMR in the natural product field (20, 21), and there is now a website devoted to this topic, which is an excellent source of information (22).

There are two ways in which quantitative ^1H NMR measurements can be carried out. The first approach, which is particularly suitable for natural product investigations, is to use these measurements to determine the relative amounts of different compounds in a complex mixture (20). There are two basic requirements. First, there must be at least one well-resolved peak (corresponding to a known number of protons) for each compound so that the relative amounts can be estimated from the integrated areas of the resolved peaks. It may be difficult to meet this requirement for complex mixtures so the use of the highest field available spectrometer is strongly recommended. Second, ideally one should know the identities of the different compounds in the mixture. Unfortunately, this often is not known and may be difficult to determine for minor components. However, if the goal is to determine the relative purity of a single major component, then the knowledge of the exact structures of minor components may not be essential (23).

Alternatively, quantitative ^1H NMR can be used to determine the absolute concentration of a compound. This requires the use of a reference standard of known concentration. In the past, some form of internal reference has been commonly used. The reference compound needs to be non-volatile (so that the amount added to the sample can be accurately weighed), not react with the compound of interest, and have a peak (preferably a singlet) in a region of the spectrum otherwise free of peaks. It is also desirable that reference compounds not have other peaks that overlap with the spectrum of the compound for which the concentration is to be

determined. The internal reference should also be a compound that can be easily separated from the natural product to facilitate recycling of rare samples after NMR spectroscopy!

An alternative approach, which is increasingly being adopted, is to use an external reference. The two major NMR spectrometer manufacturers favor different methods of referencing. The first approach, which goes by the name, ERETIC, is most commonly used with Bruker spectrometers (24). This involves electronically adding a reference signal of known intensity. The signal is injected into an unused coil on the spectrometer (*e.g.* the heteronuclear coil when acquiring proton spectra) at a clear region of the spectrum. The reference signal can be calibrated by comparing its integral with that of a standard solution of known concentration. The main problem with this approach is that the relative areas of the reference signal and the sample signals (which are detected on different coils of the probe) are slightly dependent on the nature of the sample solution, with the biggest problems occurring with “lossy” samples, particularly those with high salt concentrations. Various procedures to correct this type of error have been suggested, *e.g.* PULCON (25) and QUANTAS (26), which can reduce the uncertainties in this type of measurement to under 1%. This requires accurate recalibration of the 90° pulse width for each sample to ensure maximum precision.

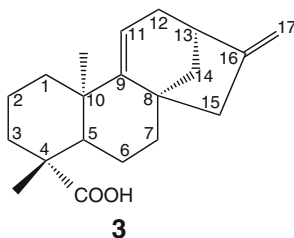
The second approach, favored by Agilent (27), takes advantage of the high linearity and reproducibility of modern NMR spectrometers. In this case, the reference standard is first measured in one tube then the sample is run in the second tube. Ideally, the same solvent should be used for both tubes. If the reference sample is measured periodically, the reported precision is similar to that reported using PULCON or QUANTAS. However, if the reference measurement is repeated before each new sample, even higher precision (well under 1%) is claimed. It appears that environmental conditions (particularly variations in the laboratory temperature) provide the largest source of error with either method and this is minimized if the calibration is repeated for each new sample (27).

¹³C NMR spectroscopy can, in principle, also be used for quantitative measurements, with the increased resolution of ¹³C spectra an attractive feature for complex mixtures. However, this is very difficult in practice. The major stumbling block is the far lower sensitivity of ¹³C NMR relative to ¹H NMR, making it very difficult to obtain accurate integrals of peak areas. In addition, there are large differences in *T*₁ relaxation times between protonated and non-protonated carbons in organic compounds, requiring extremely long relaxation delays in order to get comparable signal intensities for the two types of carbons. A related problem is that the nuclear *Overhauser* enhancements (NOEs) are often significantly smaller for non-protonated than protonated carbons. In order to make the entire spectrum quantitative, one must combine a long relaxation delay (ideally 5*T*₁ of the slowest relaxing carbons) with NOE suppression. The latter is accomplished by gating the decoupler off during the relaxation delay, turning it on only during the acquisition period. Unfortunately, this aggravates the problem of low ¹³C sensitivity by further reducing the signal/noise (*S/N*) through NOE suppression.

However, if the goal is specifically to determine the relative amounts of two or more compounds in a mixture or the absolute amount of a single component relative to a reference compound, these problems can be minimized by solely measuring peak areas for one or more protonated carbons. Particularly, T_1 values for methine or methylene carbons of natural products are typically less than 1 s, allowing the use of a relatively short relaxation delay. However, if precise quantitative data are required, NOE suppression should still be used because the NOE values for different carbons may not be exactly the same. Finally, the possibility of carrying out quantitative ^{13}C NMR measurements is significantly improved if one has access to a cryogenically cooled probe, particularly when optimized for ^{13}C detection.

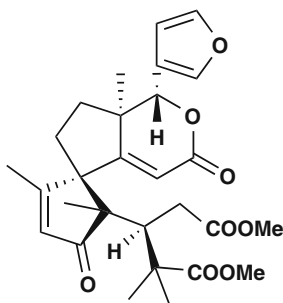
4 Using 2D NMR to Determine Skeletal Structures of Natural Products

By 1984, it was apparent that 2D NMR was a potentially powerful technique for investigating structures of natural products (28). However, the investigations up to that time had involved the use of ^1H - ^1H homonuclear correlation spectroscopy (COSY) and one-bond ^1H - ^{13}C heteronuclear correlation spectroscopy. These two types of spectra together could determine protonated carbon fragments of molecules but could not provide total structures. However, three publications in 1984 demonstrated the use of long-range (*i.e.* separated by two or three bonds) ^1H - ^{13}C heteronuclear correlation spectroscopy to determine complete structures by combining different protonated carbon fragments together into complete structures *via* correlations to quaternary carbons and/or through heteroatoms from one protonated carbon fragment to another (29–31). The first two of these publications mainly focused on correlations to carbonyl groups in polypeptides, but the third applied the technique more broadly to assign the structure and spectra of a diterpene, kauradienoic acid (**3**).



This provided the basic combination of techniques, which is still used for natural product structure elucidation, and a large number of publications using long-range correlation soon appeared. We believe that the earliest example, which most clearly indicated of the power of this technique, was provided by the determination of the structure of guyanin (**4**), a tetranortriterpene of unprecedented structure, solely by this combination of 2D experiments (32). Guyanin (**4**) has 36 heavy atoms (28 carbons and 8 oxygens) but only 17 protonated carbons and no sequences of protonated

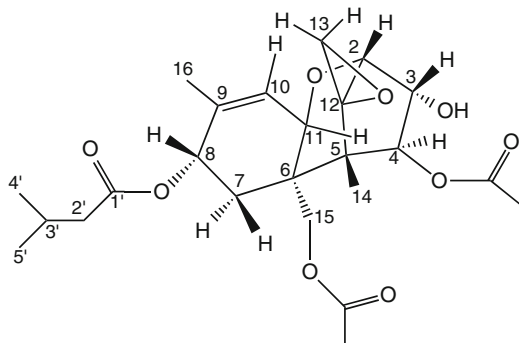
carbons greater than two. Thus, it would have been impossible to determine the full skeletal structure without long-range ^1H - ^{13}C correlation data.



4

The basic COSY sequence (33) is still in wide use to date, usually with the use of z -axis gradients for coherence pathway selection. This is sometimes supplemented by TOCSY (“total correlation”) spectra, which can provide correlations in a whole sequence of coupled protons (34). Early ^1H - ^{13}C correlation experiments involved ^{13}C (“direct”) detection. One-bond correlation spectra were mainly obtained with the HETCOR sequence (35) while long-range experiments were either obtained with the basic HETCOR sequence optimized for long-range coupling constants (29, 31) or with one of three sequences specifically designed for this purpose (30, 36). However, the ^1H - ^{13}C correlation experiments are now almost exclusively obtained using ^1H (“indirect”) sequences to take advantage of their higher sensitivity. One-bond spectra were originally mainly obtained with the HMQC sequence (37) but now often with the HSQC sequence (38). Long-range ^1H - ^{13}C correlation spectra are almost exclusively obtained with one of the variants of the basic HMBC sequence (39). The relative advantages and disadvantages of some of the alternative versions of the various 2D sequences are discussed in the following section.

The basic approach to assembling a structure from the correlation data from different 2D experiments is similar to putting together the pieces of a jigsaw puzzle. This has been illustrated in detail for T-2 toxin (5) (40). We will illustrate the same approach using spectra for kauradienoic acid (3) in Sect. 7.



5

The largest obstacle until now in using this approach to determine skeletal structures of natural products and other organic compounds is the lack of any sequence that could clearly distinguish between 2-bond and 3-bond ^1H - ^{13}C correlations to non-protonated carbons. This can lead to ambiguities and possible alternative structures. However, it has recently been shown that the 1,1-ADEQUATE sequence (41) can be used to specifically identify all two-bond H-C correlations (42, 43). In turn, this can dramatically decrease the amount of time needed and the number of alternative structures generated when using computer-assisted structure elucidation (CASE, see Sect. 8) (43). The problem is that this experiment requires one-bond ^{13}C - ^{13}C coupling, a 0.01% probability. Consequently, this is really only feasible if one has access to a ^1H -optimized cryogenically cooled probe. For example, it has recently been shown that one can obtain an acceptable quality 1,1-ADEQUATE spectrum overnight with less than 1 mg of strychnine, using a 1.7-mm cryogenically cooled 600-MHz probe (44). These authors used covariance processing of the 1,1-ADEQUATE spectrum and an edited HSQC spectrum to generate an improved quality spectrum (44). However, it should be noted that this approach still requires sufficient S/N of the original 1,1-ADEQUATE correlations for them to appear in the covariance spectrum. Covariance processing mathematically combines the results of two different 2D spectra to produce a new spectrum, which, although it does not actually contain new information, may display key correlation data in a manner that is more obvious and easier to interpret (45). It is particularly valuable when it can use two high-sensitivity spectra to generate a spectrum with good S/N that corresponds to one of much lower intrinsic sensitivity and which otherwise would take far longer to acquire.

While ^{14}N is a reasonably high sensitivity NMR nucleus (with almost 100% abundance), it is a quadrupolar nucleus, and its resultant extremely broad lines make it of extremely limited value for natural product investigations. However, ^{15}N NMR spectra can provide very useful structural information for nitrogen-containing natural products (46, 47). The low natural abundance of ^{15}N (0.37%) combined with its low frequency (~ 50 MHz on a 500-MHz spectrometer) make it unsuitable for direct detection. However, indirect (proton) detection provides a tenfold enhancement, and all recent applications in the natural product field have used indirect detection methods, mainly involving HSQC, HMQC, and HMBC spectra or variants of these experiments. There are two further problems with ^{15}N for natural product research in addition to sensitivity limitations (46). First, while one-bond ^1H - ^{15}N coupling constants fall in a narrow (90–100 Hz) range, two-bond and longer range couplings tend to show much greater variability than the corresponding ^1H - ^{13}C couplings (46). This makes it difficult to detect all expected long-range ^1H - ^{15}N correlations in an HMBC experiment with a fixed ^1H - ^{15}N delay. Second, ^{15}N has an extremely large chemical shift range (*ca.* 600 ppm) while most H/X and H/C/N probes have relatively long ^{15}N 90° pulse widths (*ca.* 40 μs), which cannot uniformly excite the entire ^{15}N spectral window, risking failure to detect peaks near either end of the window.

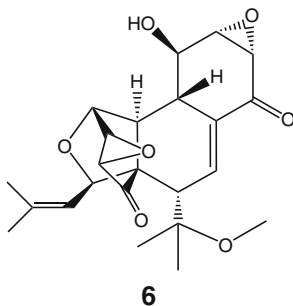
Fortunately, advances in instrumentation and pulse sequence design have minimized these problems. First, cryogenically cooled probes have significantly reduced

the sample requirements. For example, *Martin et al.* have shown that a recently developed 1.7-mm cryo-microprobe can provide useable ^{15}N HMBC spectra overnight with well under 1 mg of sample (48). The same probe had a 90° pulse width of 25 μs , allowing 90% or more excitation over about a 500-ppm shift range for ^{15}N (48). Various solutions have been proposed to minimize the problem of the variation in long-range ^1H - ^{15}N coupling constants (47). As an example, *Cheatham et al.* have developed the ^{15}N CIGAR sequence, which uses an “accordion” delay that is optimized for ^1H - ^{15}N couplings in the 3–10 Hz range (49). There are at least two other time-saving approaches. First, since there are typically only a very small number of ^{15}N peaks in a nitrogen-containing natural product (often only one), one can usually utilize a lower $F1$ resolution by decreasing the number of time increments used (and correspondingly increasing the number of scans per time increment to improve sensitivity).

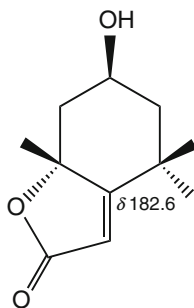
Second, if the natural product contains both protonated and non-protonated nitrogens, one does not need to acquire separate one-bond (HSQC or HMQC) spectra along with an HMBC-type spectrum. Instead, one can eliminate the “ J filter” present in HMBC sequences and simultaneously observe both one-bond and n-bond correlations, with the former distinguished from the latter by the observation of a large (90–100 Hz) doublet splitting due to the direct ^1H - ^{15}N coupling (49).

5 Avoiding Getting the Wrong Skeletal Structure

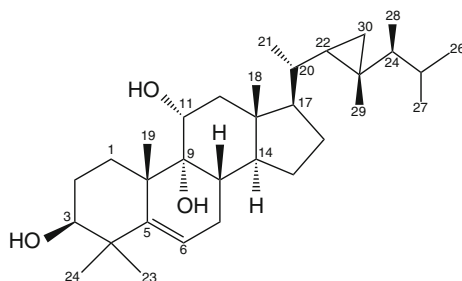
One might assume that, with all of the modern techniques available, it would be unlikely that incorrect structures are reported in the chemical literature. However, there have been a surprisingly large number of incorrect structures reported (50). Many of these involve errors in stereochemistry, which often can be quite tricky to determine, particularly in acyclic compounds or those with conformationally mobile rings (see Sect. 6 for a discussion of the determination of stereochemistry). However, some involve incorrect skeletal structures where errors should be easier to avoid. One example of the latter kind of error, which attracted a lot of attention and controversy, is hexacyclinol (**6**) (51). Fortunately, there are several precautions that one can take to minimize the risk of mistakes. These are discussed below.



1. Do not try to fit the data to a preconceived notion of structure but rather allow the data to determine possible structure(s). This problem is most likely to occur if there is severe overlap in at least part of the spectrum and/or marginal signal/noise, which may cause ambiguities in assigning correlations. In this situation, trying to save time by collecting too few scans and/or too few data points (particularly $F1$ increments) is actually counterproductive by maximizing the risks of ambiguity. An old trick of aromatic solvent-induced shifts (52) can often be used to minimize overlap by repeating the spectrum with added increments of C_6D_6 . The present authors have found this to often be valuable for natural product structure elucidation (13).
2. Tabulate all 2D data and check carefully for unexpected peaks and missing expected peaks. A peak of significant intensity in an HMBC spectrum, which is not consistent with a proposed structure, should be taken as a strong warning sign that the structure is probably incorrect. On the other hand, the absence of an expected peak may not always be as significant since this may just be due to a relatively small 2-bond or 3-bond 1H - ^{13}C coupling constant. A common case where a peak is either not observed or very weak is for 2-bond correlations in aromatic or olefinic groups. Another case where weak peaks often occur is for 3-bond correlations involving *axial* protons in cyclohexane-like rings. A general knowledge of expected n -bond 1H - ^{13}C coupling constants for different types of structural units is helpful (53). Some representative values for aliphatic, olefinic, and aromatic derivatives are: 0–3 Hz for 2-bond olefinic and aromatic couplings, 8 Hz for 3-bond aromatic and *cis*-olefinic couplings, 12 Hz for *trans*-olefinic couplings, 3–5 Hz for 2-bond aliphatic couplings, 2–4 Hz for 3-bond *gauche*-aliphatic couplings and 6–9 Hz for *anti*-aliphatic couplings. Longer range (4-bond and 5-bond) couplings are generally less than 2 Hz but may occasionally show up as weak HMBC correlations, particularly in conjugated derivatives or from methyl protons.
3. Beware of deceptively appearing spectra. In our experience this may take one of two forms. First, due to a combination of effects, a ^{13}C peak may occur at a chemical shift, which seems more consistent with an entirely different type of functional group. For example, a conjugated lactone (7) had a ^{13}C peak at δ 182.6 ppm that was initially assumed, based on the chemical shift, to be either a carboxylic acid group or a highly conjugated ketone. However, it was eventually assigned as the non-protonated olefinic carbon, based on HMBC cross-peaks to all three methyl proton signals and the two pseudo-equatorial methylene protons (54). A second situation is when there is accidental equivalence between two coupled proton signals. While one learns very early in NMR that no coupling will be observed between equivalent protons, it is still easy to forget this in the case of accidental equivalence. One example of this was provided by the marine sterol gorgost-5-ene-3 β ,9 α ,11 α -triol (8) (55). In this case, the expected H-21 methyl doublet appeared as a slightly broadened singlet, initially leading to the assumption of some type of rearrangement of the steroid side chain. However, examination of the HSQC spectrum showed that H-20 and H-21 had the same chemical shift while the HMBC spectrum showed a strong cross-peak between H-21 and C-20, confirming that the methyl singlet was due to accidental equivalence of H-20 and H-21.
4. Just because a structure appears to fit the correlation data, do not assume that this is the only structure consistent with the data. This error, which is extremely easy



7



8

to make, can be best avoided by the use of computer-assisted structure elucidation (CASE) (see Sect. 8). The program Structure Elucidator (56), provided by ACD, is the one with we have the most experience. This will not only determine all possible structures consistent with the correlation data but will rank them in order of probability based on a comparison of observed ^1H and ^{13}C chemical shifts with those calculated by the program for the different structures. In cases where there is no clear distinction between two or more possible skeletal structures, examination of the structure may suggest additional experiments, which may allow this distinction (57). The program also alerts the user to ambiguous assignments due to severe spectral overlap so that these problems can be considered individually (57).

6 Determination of Configuration and/or Conformation of Natural Products

The degree of difficulty of this task is mainly determined by whether the molecule exists in a single, relatively rigid, conformation or is undergoing rapid interconversion between two or more conformations. The latter situation is particularly difficult to deal with, but even determining the configuration and conformation of a rigid molecule can still pose problems. Also note that NMR alone will rarely provide absolute configurations. We will begin by considering the investigation of rigid molecules.

The two main tools to address these problems are *vicinal* ^1H - ^1H coupling constants (58) and nuclear *Overhauser* enhancements (NOEs) (59). *Vicinal* coupling constants can be used to estimate ^1H -C-C- ^1H dihedral angles, usually with the aid of some type of *Karplus* relationship such as the *Altona* equation (60). However, one must recognize that splittings in a proton multiplet are not always identical to coupling constants, particularly in the case of strong coupling. For example, if a pair of methylene protons has a near-zero chemical shift difference, they will appear to be equally coupled to an adjacent methine proton, regardless of the actual *vicinal* coupling constants, a phenomenon known as “virtual coupling” (61). However, provided that the methylene protons are, in principle, diastereotopic, it may be possible to separate them by using solvent effects such as aromatic solvent-induced shifts and thus determine the *vicinal* couplings. Another problem occurs if the two *vicinal* protons are equivalent by symmetry or by accident since one then will not observe a coupling between them. In either case, a coupled HSQC spectrum will allow one to measure the coupling since the large one-bond CH coupling within the ^1H - ^{13}C - ^{12}C - ^1H fragment effectively makes the protons non-equivalent (62, 63). This is a modification of the old idea of using ^{13}C satellites in a proton spectrum for this purpose (64). Alternatively, in the case of accidental equivalence, one could again try to use solvent effects to separate the two proton signals.

Nuclear *Overhauser* enhancements can be measured either from 2D NOESY (65) or ROESY (66) spectra or by selective 1D equivalents. These have usually been used in a qualitative “yes–no” sense to determine whether pairs of protons are relatively close or relatively far apart. However, driven by improvements in selective 1D pulse sequences (67) and spectrometer hardware, there has been a recent revival of interest in using selective 1D NOESY or ROESY measurements to make quantitative distant measurements in organic molecules (68). Unlike, the earlier NOE-difference measurements which measure steady-state NOEs, the selective 1D experiments (and their 2D analogues) measure transient NOEs by selectively inverting a chosen proton multiplet and then following the buildup of NOEs for other cross-relaxed protons as the inverted proton relaxes back towards its initial value (59). The usual approach has been to plot NOE intensity *versus* mixing time and determine the slope from the initial linear region of the plot, the initial rate approximation (IRA) (59). However, this can be time consuming. Recently, *Butts et al.* have championed the alternative of using a single mixing time, which is short enough to fall in the regime where the IRA holds (68). By using the known distance between methylene protons as a calibration point, this group has not only measured the distances between other pairs of protons in strychnine with surprising accuracy (~3%) but also detected the presence of a previously unknown minor conformation of strychnine (68). The main limitation of this approach is that for higher molecular weight compounds and/or viscous solutions, the IRA may only be valid for short mixing times (<0.2 s) where the NOE peaks will be quite weak. However, the range of acceptable mixing times can be significantly extended by applying the PANIC correction developed by *Macura* for 2D NOESY measurements (69) and extended by *Hu* for 1D NOESY and ROESY measurements (70). This requires a second measurement for each proton with zero mixing time with a correction for the NOE

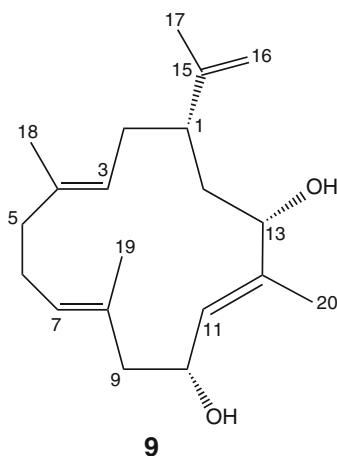
peaks obtained from the ratio of the areas of the inverted peak with zero and the chosen mixing time.

An alternative approach, which is being increasingly applied, is to measure residual dipolar ^1H - ^{13}C and/or ^1H - ^1H couplings for a molecule in a weakly aligning medium (71, 72). The relative magnitudes of different residual dipolar couplings are determined by the relative orientations of C-H or H-H bond vectors relative to the alignment tensor of the molecule. Unfortunately, this has a $(3\cos^2\theta - 1)$ relationship, and thus one cannot determine directly whether an individual bond has an angle θ , or $180^\circ - \theta$, with respect to the alignment tensor. Nevertheless, provided one can measure at least five residual dipolar couplings, one can determine the alignment tensor (71). Obviously, the more values that can be determined, the better the molecular configuration and conformation will be defined. A series of alignment media suitable for polar and non-polar organic solvents are available, and measurements of 1-bond ^1H - ^{13}C and *geminal* ^1H - ^1H dipolar couplings are usually carried out with one or more modified versions of the coupled HSQC sequence (71). This is performed by determining the differences of these couplings in isotropic and partially aligning media. A number of relative configurations and conformations of rigid organic molecules have been determined in this way (71, 72). One recent intriguing example of the power of this technique was the determination of the structure of a reaction impurity, which could not be determined by 2D NMR techniques (73). Nevertheless, this is a laborious technique, which, in our opinion, should be regarded as a last-resort method.

Any of the three approaches discussed above can be carried out in conjunction with calculations of molecular conformation of rigid molecules, either with classical molecular mechanics calculations or more increasingly with *ab initio* quantum mechanical calculations. However, calculations to determine the relative free energies and consequently the relative populations of different conformations are absolutely essential if one is dealing with a flexible molecule with two or more significantly populated conformations. Unfortunately, even small differences in calculated free energies correspond to significant differences in populations. Thus, while there has been some progress in interpreting both NOE measurements (74) and residual dipolar couplings (71) for flexible molecules, the conversion of the weight-averaged data from these measurements into accurately determined conformations of these molecules will require increasingly accurate energy calculations to make this at least semi-routine. Fortunately, the increasing speed of computers may make this feasible in the near future.

Normally, one can only determine the relative configuration of a natural product from NMR data. However, the conversion of secondary CH(OH) groups into chiral esters by Mosher's method can often determine the absolute stereochemistry of that carbon in a natural product (75). While this is an empirical approach based on induced chemical shifts by the chiral ester, it seems surprisingly reliable. One example where we found this to be valuable was in determining the absolute stereochemistries of the two CH(OH) centers of the cembrane, cleospinol-A (**9**), which contained a flexible 14-membered macrocyclic ring (76). With the knowledge of the absolute stereochemistry at these two centers and the configurations of other chiral

centers relative to these two, it was possible to determine the absolute stereochemistry of the entire molecule. Another approach is to combine NMR measurements of relative stereochemistries within a molecule with a chiral spectroscopic method such as vibrational circular dichroism (see this volume: *Joseph-Nathan P and Gordillo-Román B* (2014) *Vibrational Circular Dichroism Absolute Configuration Determination of Natural Products*. *Progr Chem Org Nat Prod* 100:311), which gives the absolute stereochemistry (77).



7 An Example of a Solved Structure: Kauradienoic Acid

A considerable number of NMR experiments have been introduced and discussed in the previous section. In order to illustrate how a series of such experiments is used in practice, the authors will take readers through a complete structural elucidation exercise for the compound kauradienoic acid (**3**). The structure of this compound was originally proposed in 1971 (78), and its ^1H and ^{13}C NMR spectra were assigned completely in 1984 (31). If one had no knowledge of the compound's structure, such a study would begin with the determination of its infrared, ^1H , and ^{13}C NMR spectra, and high-resolution mass spectrum. This compound is a white solid that is soluble in chloroform.

The mass spectrum indicated a molecular weight of *ca.* 300 and provided several pieces of information about the compound. The even molecular weight establishes that the unknown contains an *even* number of nitrogen atoms and probably none (zero, as in spectral-edited experiments, is considered to be an even number). If we make the preliminary, and usually justified, assumption that the unknown compound contains only carbon, hydrogen, and oxygen, the high-resolution, mass spectrometric-determined molecular weight supports a molecular formula of $\text{C}_{20}\text{H}_{28}\text{O}_2$. This formula, in turn, dictates the presence of seven units of unsaturation.

A strong infrared absorption centered at $1,760\text{ cm}^{-1}$ suggests the presence of a carbonyl group. Moreover, the occurrence of two oxygens in the molecular formula supports the inference that the carbonyl absorption could be due to an acid or ester group.

7.1 HSQC Data

After initial ^1H and ^{13}C NMR spectra have been determined for an unknown compound, it is useful to establish which hydrogens are directly attached to specific carbons by an HMQC or HSQC experiment. The latter is the experiment of choice because of its better resolution in the $F1$ domain (see Sect. 10). An edited-HSQC spectrum of a small sample of the unknown compound is shown in Fig. 1. Coupling constants, which have been measured in the ^1H NMR spectrum for those multiplets that are not severely overlapped, correspond to proton cross peaks in the COSY spectrum (see Sect. 7.2) and are included in parentheses in Table 1.

A cursory examination of the HSQC data in Table 1 indicates that the unknown compound is essentially aliphatic in nature. The highest frequency signal (182.74 ppm) is in the acid carbonyl range, while those carbons with chemical shifts from 158.56 to 105.48 ppm appear to be alkenic and require the presence of one methylene (105.48 ppm), one methine (114.90 ppm), and two quaternary sp^2

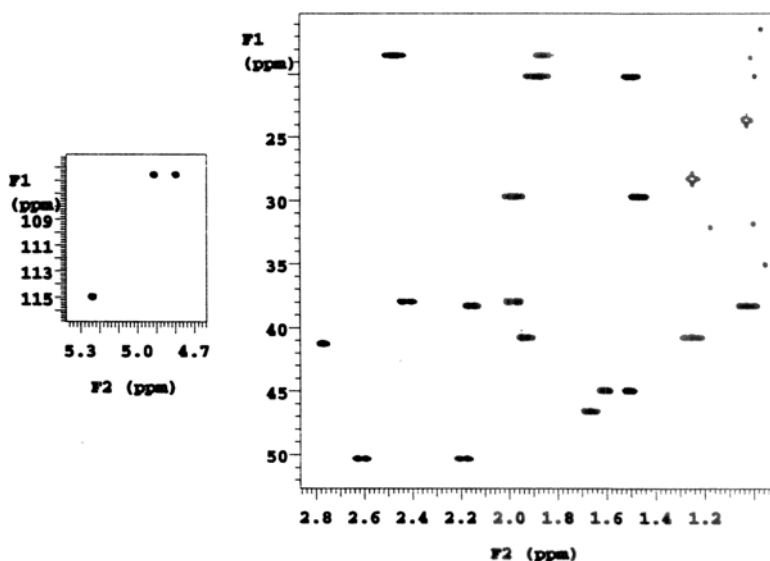


Fig. 1 Gradient-selected HSQC spectrum for kauradienoic acid (**3**) with aliphatic region on the right and olefinic region on the left

Table 1 HSQC data for kauradienoic acid (**3**)

δ_c /ppm	Attached hydrogens (with selected couplings)	
182.74	–	
158.56	–	
155.92	–	
114.90	5.24 ddd (4, 2.7, <1)	
105.48	4.91 brs	4.80 brs
50.32	2.62	2.20
46.56	1.67 d (11)	
44.94	1.60 d (5.2)	1.50 d (2.7)
44.69	–	
42.27	–	
41.24	2.77 dddd (5.2, 4.6, 2.7, 2, <1)	
40.75	1.93	1.24
38.79	–	
38.31	2.16	1.02
37.93	2.43 dd (4.6, 2.7)	1.99 dd (4, 2)
29.66	1.97 dd (10.8, 9)	1.46 dd (9, 1)
28.23	1.24 s (CH ₃)	
23.62	1.02 s (CH ₃)	
20.16	1.88	1.50
18.48	2.47 dddd (11, 9, 8.7, 1)	1.68 ddd (11, 8.9, 8.7)

carbons. Resonances of the remaining carbons are generally in the chemical shift range for aliphatic carbons that are not attached to oxygen.

The HSQC spectrum reveals the presence of two methyl, nine methylene, and three methine carbons. Subtracting these 14 protonated carbons from the 20-carbon total identifies the remaining six quaternary carbons. In addition, summing the six methyl, 18 methylene, and three methine protons accounts for 27 hydrogens. Since the molecular formula requires 28 hydrogens, the final proton must be attached to oxygen.

7.2 COSY and TOCSY Data

The construction of proton spin-coupling networks is achieved by COSY and TOCSY experiments. ¹H spin systems are observed in the COSY (Fig. 2) and TOCSY (Fig. 3) contour plots. Two are relatively simple and comprise (i) a 6-spin system of three contiguous methylene groups: a terminal pair of protons (1.93 and 1.24 ppm), an interior pair (1.88 and 1.50 ppm), and another terminal pair (2.16 and 1.02 ppm) (**10**) and (ii) a 5-spin system containing an isolated proton (1.67 ppm), an adjacent (middle) methylene pair of protons (2.47 and 1.68 ppm) and another methylene pair (1.97 and 1.46 ppm) (**11**).

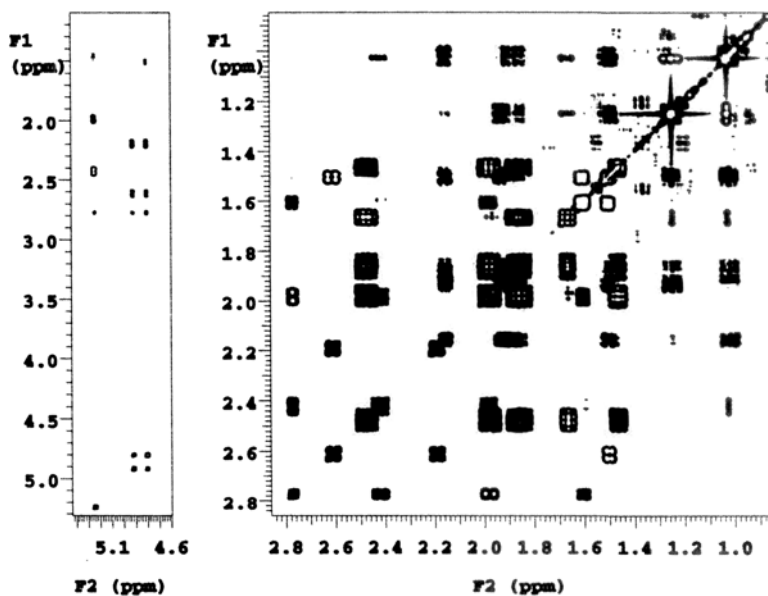


Fig. 2 Gradient-selected absolute-value COSY spectrum of kauradienoic acid (3). The *spectrum on the right* shows correlations between aliphatic protons while that on the *left* shows correlations between olefinic and aliphatic protons

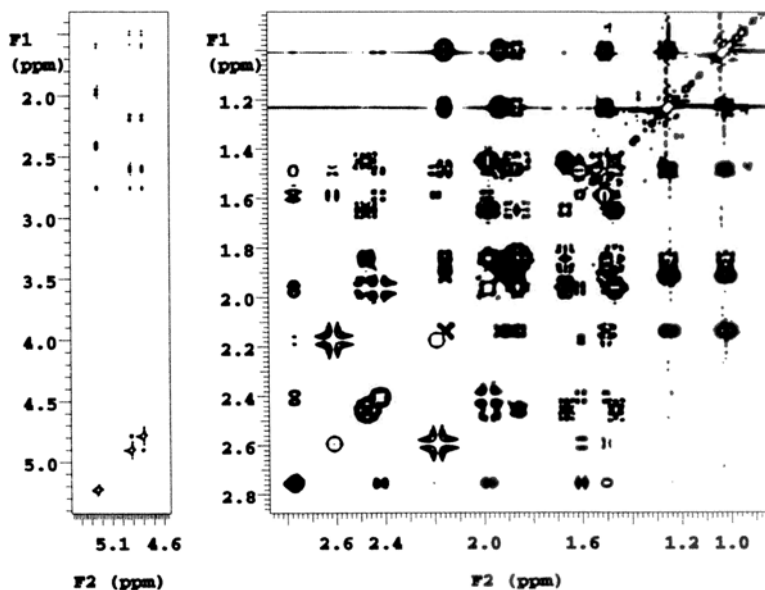
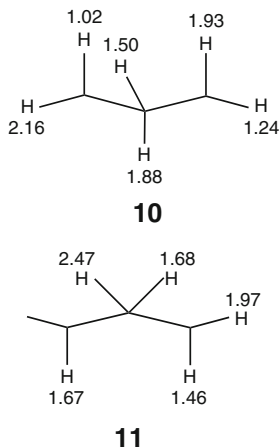


Fig. 3 Gradient-selected TOCSY spectrum of kauradienoic acid (3). The expansions are similar to those in Fig. 2. The spectrum was obtained with the original TOCSY sequence, without a zero-quantum filter, and shows some distortions of multiplet structures



The third spin system is extensive and encompasses ten protons: two methines (5.24 and 2.77 ppm) and four methylenes (4.91 and 4.80, 2.62 and 2.20, 2.43 and 1.99, and 1.60 and 1.50 ppm). A 1D-Z-TOCSY trace, through the alkenic proton (Fig. 4), elegantly illustrates how it is coupled to the methylene protons at 2.43 and

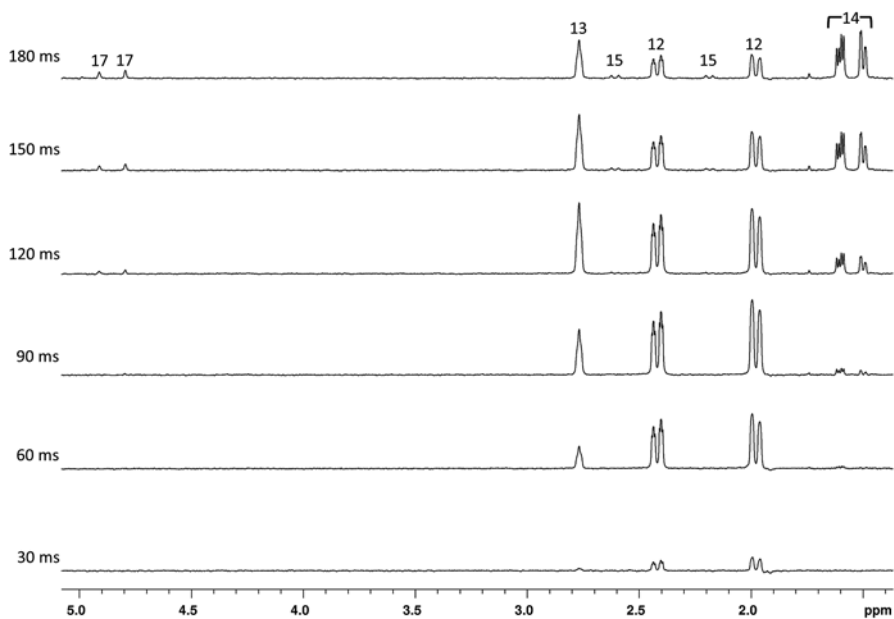


Fig. 4 1D Z-TOCSY spectra with different mixing times, obtained by selective irradiation of H-11 (5.24 ppm) of kauradienoic acid (**3**). Mixing times are listed at the *left* of each spectrum and proton assignments are listed at the *top*. The spectra illustrate nicely how a sequence of coupled protons can be fully assigned by a series of 1D TOCSY spectra with incremented mixing times. Also note that the zero-quantum filter allows observation of undistorted proton multiplets

1.99 ppm, the methine proton at 2.77 ppm, the methylene protons at 1.60 and 1.50 ppm, and finally to the alkenic protons at 4.91 and 4.80 ppm and the methylene protons at 2.62 and 2.20 ppm (**12**). The data for these spin systems are summarized in Tables 2 and 3.

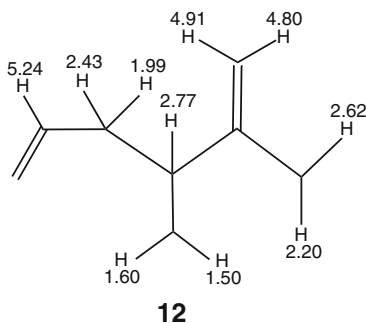


Table 2 COSY data for kauradienoic acid (**3**)

δ_{H} /ppm	Spin-coupled partners
5.24	2.77, 2.43, 1.99
4.91	4.80, 2.77, 2.62, 2.20
4.80	4.91, 2.77, 2.62, 2.20
2.77	5.24, 4.91, 4.80, 2.62, 2.43, 1.99, 1.60, 1.50
2.62	4.91, 4.80, 2.77, 2.20, 1.50
2.47	1.97, 1.68, 1.67, 1.46
2.43	5.24, 2.77, 1.99
2.20	4.91, 4.80, 2.62
2.16	1.88, 1.50, 1.02
1.99	5.24, 2.77, 2.43, 1.60
1.97	2.47, 1.68, 1.46
1.93	1.88, 1.50, 1.24
1.88	2.16, 1.93, 1.50, 1.24, 1.02
1.68	2.47, 1.97, 1.46
1.67	2.47
1.60	2.77, 1.99, 1.50
1.50 ^a	2.77, 2.62, 2.16, 1.93, 1.88, 1.60, 1.24, 1.02
1.50 ^a	
1.46	2.47, 1.97, 1.68
1.24	1.93, 1.88, 1.50
1.24 (CH ₃)	(none) ^b
1.02	2.16, 1.88, 1.50
1.02 (CH ₃)	(none) ^c

^aCorrelations cannot be differentiated for the pair of hydrogens at 1.50 ppm because of their identical chemical shift

^{b,c}Since the methyl signals are singlets, COSY correlations must be to the individual hydrogens at 1.24 and 1.02 ppm, respectively

Table 3 TOCSY data for kauradienoic acid (**3**)

$\delta_{\text{H}}/\text{ppm}$	Spin-system members
5.24	2.77, 2.43, 1.99, 1.60, 1.50
4.91	4.80, 2.77, 2.62, 2.43, 2.20, 1.99, 1.60, 1.50
4.80	4.91, 2.77, 2.62, 2.43, 2.20, 1.99, 1.60, 1.50
2.77	5.24, 4.91, 4.80, 2.62, 2.43, 2.20, 1.99, 1.60, 1.50
2.62	4.91, 4.80, 2.77, 2.20, 1.99, 1.60, 1.50
2.47	1.97, 1.68, 1.67, 1.46
2.43	5.24, 2.77, 1.99, 1.60, 1.50
2.20	4.91, 4.80, 2.77, 2.62, 1.60, 1.50
2.16	1.93, 1.88, 1.50, 1.24, 1.02
1.99	5.24, 2.77, 2.62, 2.43, 1.60, 1.50
1.97	2.47, 1.68, 1.67, 1.46
1.93	2.16, 1.88, 1.50, 1.24, 1.02
1.88	2.16, 1.93, 1.50, 1.24, 1.02
1.68	2.47, 1.97, 1.46
1.67	2.47, 1.97, 1.46
1.60	5.24, 4.91, 4.80, 2.77, 2.62, 2.43, 2.20, 1.99, 1.50
1.50 ^a	5.24, 4.91, 4.80, 2.77, 2.62, 2.43, 2.20, 2.16, 1.99
1.50 ^a	1.93, 1.88, 1.60, 1.24, 1.02
1.46	2.47, 1.97, 1.68, 1.67
1.24 ^b	2.16, 1.93, 1.88, 1.50, 1.02
1.24 (CH ₃)	(none)
1.02 ^c	2.16, 1.93, 1.88, 1.50, 1.24
1.02 (CH ₃)	(none)

^aTOCSY correlations cannot be differentiated for the pairs of hydrogens at 1.50, 1.24, and 1.02 ppm because of their identical chemical shifts

^{b,c}Since the methyl signals are singlets, TOCSY correlations must be to the individual hydrogens at 1.24 and 1.02 ppm, respectively

7.3 HMBC Data

Since the ¹³C NMR spectrum of the unknown compound is far less congested than its ¹H spectrum, HMBC is the experiment of choice to establish the longer range, C-H correlation networks. HMBC and CIGAR-HMBC spectra of the unknown compound were recorded, and the HMBC spectrum is shown in Fig. 5. The HMBC data from these contour plots are summarized in Table 4 and will be extensively discussed below in Sect. 7.5.

Two points that can now be made are that, first, methyl signals are particularly helpful in the interpretation of HMBC spectra. Just as their signals are almost

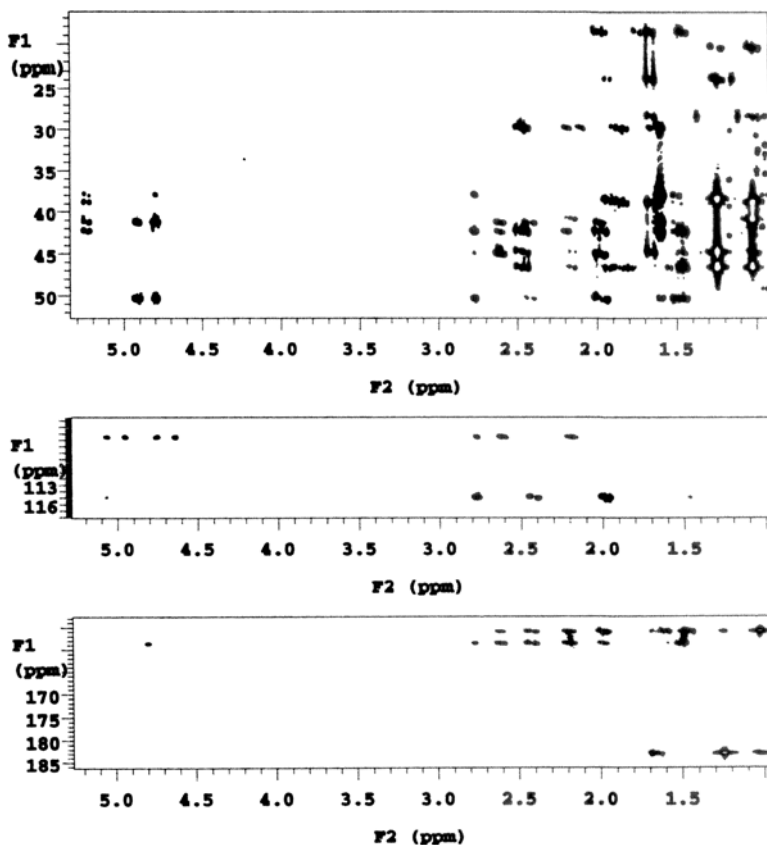


Fig. 5 Expansions of a gradient-selected absolute-value HMBC spectrum for kauradienoic acid (3), illustrating the 2-bond and 3-bond C–H correlations for different ¹³C regions

always the largest resonances in ¹H NMR spectra, so their cross-peaks are generally the strongest in HMBC spectra where they exhibit prominent 2-bond and 3-bond connectivities. Second, in cyclohexane chair conformers, *vicinal* C–H couplings due to *equatorial* protons are generally larger to ring carbons three bonds removed than those due to *axial* protons because of the *ca.* 180° dihedral angles of the former compared to the *ca.* 60° dihedral angles of the latter. As a result, *equatorial* protons generally give rise to stronger three-bond HMBC cross-peaks to ring carbons than their *axial* counterparts.

Table 4 HMBC data^a for kauradienoic acid (**3**)

δ_c /ppm	Correlated hydrogen chemical shifts
182.74	2.16, 1.67(s), 1.24, 1.02(s)
158.56	5.24, 2.62, 2.43, 2.20(s), 1.99, 1.60(s), 1.50, 1.02
155.92	2.62, 2.43, 2.20, 1.99, 1.60, 1.50
114.90	2.77, 2.43, 1.99
105.48	2.77, 2.62, 2.20
50.32	5.24, 4.91, 4.80, 2.77, 1.97, 1.60, 1.50
46.56	2.47, 2.16, 1.93, 1.46, 1.24, 1.02
44.94	2.62, 2.43, 1.99, 1.97, 1.46
44.69	2.47, 2.16, 1.68, 1.67, 1.50(s), 1.24, 1.02
42.27	5.24, 2.77, 2.62, 2.47, 2.20, 1.97, 1.60, 1.50, 1.46
41.24	5.24, 4.91, 4.80, 2.62, 2.43, 1.99, 1.60, 1.50
40.75	2.16(s), 1.88, 1.67, 1.50, 1.02
38.79	5.24, 1.93, 1.88, 1.68, 1.67, 1.50(s), 1.24, 1.02
38.31	1.93(s), 1.88, 1.50, 1.24
37.93	5.24, 4.80, 2.77, 1.60, 1.50
29.66	2.47, 2.20, 1.68, 1.67, 1.60
28.23	1.67, 1.02
23.62	1.93, 1.67(s), 1.24(s)
20.16	2.16, 1.93, 1.24, 1.02
18.48	1.97, 1.67, 1.46

^aFrom HMBC and CIGAR-HMBC experiments

7.4 General Molecular Assembly Strategy

Of the many techniques available to the NMR spectroscopist in structural elucidations, none are so valuable as the indirect, chemical shift correlation experiments like HMBC, TOCSY (both homo- and heteronuclear varieties), and FLOCK (36b). FLOCK is an X-nucleus detected experiment analogous to HMBC. While it is significantly less sensitive than HMBC, it is useful in those instances where very high ¹³C resolution is essential. Once molecular fragments have been identified by the COSY and HSQC experiments, combination of these fragments is attempted by means of the above techniques. As indispensable as these methods have become to NMR spectroscopists, they suffer a common limitation in that two-bond, C–H couplings cannot generally be directly distinguished from three-bond, C–H coupling constants. However, these two classes of C–H couplings can be differentiated, for adjacent protonated carbons, by examination of HSQC and COSY spectra (*vide infra*).

The process of molecular assembly can be approached in the following manner. If possible, a carbon atom is selected from which the remainder of the molecular skeleton can be built in just *one* direction. Methyl groups are, of course, excellent starting points. As mentioned above, adjacent protons, if any, can be identified from a COSY contour plot and longer-range coupled protons from a TOCSY spectrum. Conversely, the carbons to which the just-identified protons are directly attached can be determined from an HSQC or HMQC plot. The HMBC or FLOCK spectra can

then be scanned using either the contour plot, for uncongested spectra, or methyl-proton or methyl-carbon traces (more commonly the latter, but the choice depends on which spectral axis is less congested) if spectral congestion or weak cross peaks are a problem. Cross-peaks may be found for (i) the adjacent carbon (if methyl-proton traces are viewed) or protons (if methyl-carbon traces are observed, again, more likely), which represent two-bond couplings in either case, and (ii) any other carbons or protons, indicative of three-bond couplings (but always being mindful that one, or more, members of the latter group could possibly be due to ${}^nJ_{\text{CH}} > 3$). The fortunate redundancy of these 2D experiments is seen, whereby an adjacent carbon may be identified by a combination of COSY (${}^3J_{\text{HH}}$) and HSQC/HMQC (${}^1J_{\text{CH}}$) connectivities and also by HMBC/FLOCK (${}^2J_{\text{CH}}$) correlations.

The third carbon atom in the fragment (two carbons removed from the original methyl group) will likely show two-bond, C–H couplings with adjacent proton(s), if present, both backward to the second carbon and forward to the fourth carbon in the series. Carbon-atom connectivities can thus be built up using (i) two- and three-bond, C–H couplings to generally confirm previously determined C–C correlations and then (ii) three-bond, H–H and C–H couplings to extend the developing molecular structure.

Even when a methyl ${}^1\text{H}$ signal is nothing more than a broadened singlet, the COSY spectrum (either the contour plot or the methyl trace) can be scanned for cross-peaks due to long-range coupling. Turning then to HMBC or FLOCK spectra, either the contour plot or the methyl ${}^{13}\text{C}$ trace (HMBC spectrum) or ${}^1\text{H}$ trace (FLOCK spectrum) can be examined for cross-peaks due to (i) two-bond coupling to adjacent (quaternary) carbons and (ii) three-bond coupling to farther-removed carbons (with the usual ${}^nJ_{\text{CH}} > 3$ caveat). Finally, a note of caution should be mentioned. Like NOEs, the intensities of three-bond, C–H correlations are not necessarily symmetrical, *e.g.* in the four-carbon fragment pictured in Fig. 6, a strong cross-peak may, in fact, be observed between H_A and C_3 while a weak one, or none at all, is seen between H_C and C_1 . The main reason for these weak or missing correlations is the dependence of vicinal couplings on the H–A–C–1–C–2–C–3 and H–C–C–3–C–2–C–1 dihedral angles, which are seldom identical. Since similar considerations are largely absent for two-bond couplings, cross-peaks should be detected between H–A and C–2 and also between H–B and C–1. Extensive redundancy of the type described above, however, is in fact observed routinely for *vicinal* C–H correlations and is invaluable in the construction of molecular structures.

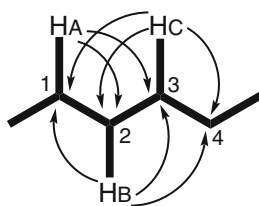
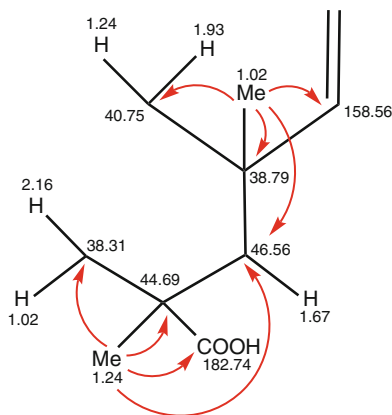


Fig. 6 Examples of 2-bond and 3-bond C–H connectivities in a typical organic chemical fragment in which ${}^3J_{\text{H}_\text{A}\text{C}_3}$ can be quite different from ${}^3J_{\text{H}_\text{C}\text{C}_1}$ and results in HMBC cross-peaks of differing intensities

The same factors that influence H–H couplings, *e.g.* dihedral angle dependence, substituent electronegativity, bond length, and bond order, apply to C–H couplings. As a general rule, C–H coupling constants are approximately 2/3 the value of the corresponding H–H couplings. In alkenes, for example, average *cis* and *trans* H–H couplings are *ca.* 11 and 18 Hz, while the corresponding C–H coupling constants are *ca.* 7 and 12 Hz.

7.5 A Specific Molecular Assembly Procedure

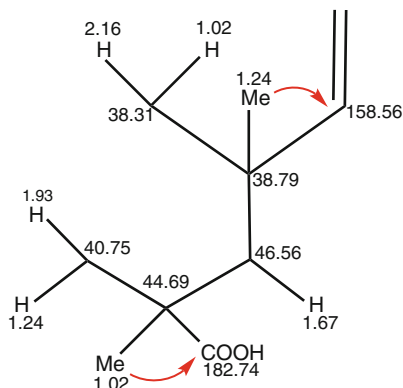
Examination of the attached-proton data in Table 1 demonstrates the presence of two methyl groups, which are both singlets at 1.24 and 1.02 ppm. Inspection of the HMBC and CIGAR-HMBC data shows that the methyl signal at 1.24 ppm exhibits strong connectivities to the carbons at 182.74, 46.56, 44.69, and 38.31 ppm. Further observation reveals that the methyl signal at 1.02 ppm displays strong connectivities to the carbons at 158.56, 46.56, 40.75, and 38.79 ppm. In addition, both methyl groups have an HMBC correlation to the carbon at 46.56 ppm. Obtaining the chemical shifts of the protons attached to these carbons from Table 1 yields the molecular fragment **13**.



13

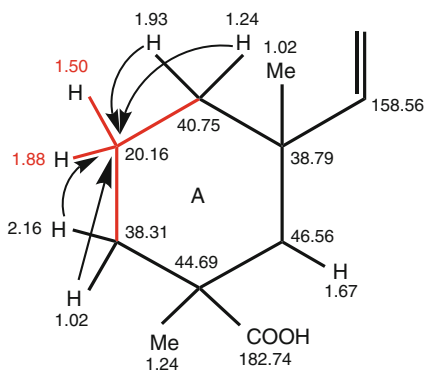
At this point, an ambiguity arises because there also happen to be two methine protons with the same chemical shifts as the two methyl groups, *viz.* 1.24 and 1.02 ppm, and in sufficiently close proximity to these two methyl groups that there is a possibility that the assignments of both the methyl groups and their adjacent methylene groups could be interchanged (**14**), a consideration that will be further examined in Sect. 8.3. The HMBC data in Table 4 show that the methyl group at 1.02 ppm in **14** could exhibit an HMBC connectivity to the carboxyl-carbon (182.74 ppm). However, a problem arises with the potential reversed assignments because the methyl group at 1.24 ppm in **14** should, likewise, display a strong HMBC connectivity to the alkenic-carbon at 158.56 ppm, but none is observed. The absence of such an important HMBC correlation, subsequent NOESY correlations, and critical HMBC traces (Sect. 7.6)

dictate that the reversed assignments given in **14** are incorrect. We will discuss this reversed-assignment situation later, at greater length, in this section.

**14**

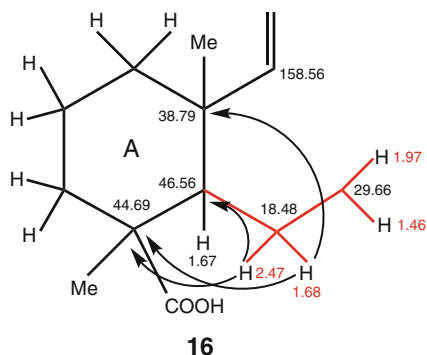
Note too that the chemical shift of the carbon at 182.74 ppm is appropriate for a carboxylic acid, and this carbon has been assigned as such. The two oxygens required by the molecular formula are thus accounted for.

Fragment **13** contains the terminal methylene groups shown in **10**, and insertion of the central methylene group of **10** to close the A-ring produces **15** (note that additions to an existing fragment are shown in red).

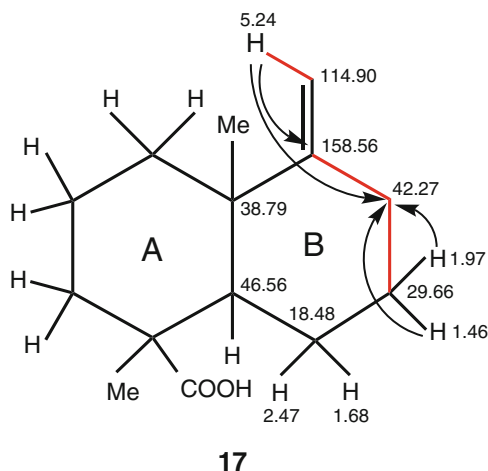
**15**

Both carbons at 40.75 and 38.31 ppm in **15** display HMBC connectivities to protons at 1.88 and 1.50 ppm. In addition, the protons at 1.93 and 1.24 ppm (on the carbon at 40.75 ppm) and the protons at 2.16 and 1.02 ppm (on the carbon at 38.31 ppm) exhibit HMBC correlations to the carbon at 20.16 ppm, to which the protons at 1.88 and 1.50 are attached. At this point, three of the seven units of unsaturation have been accounted for.

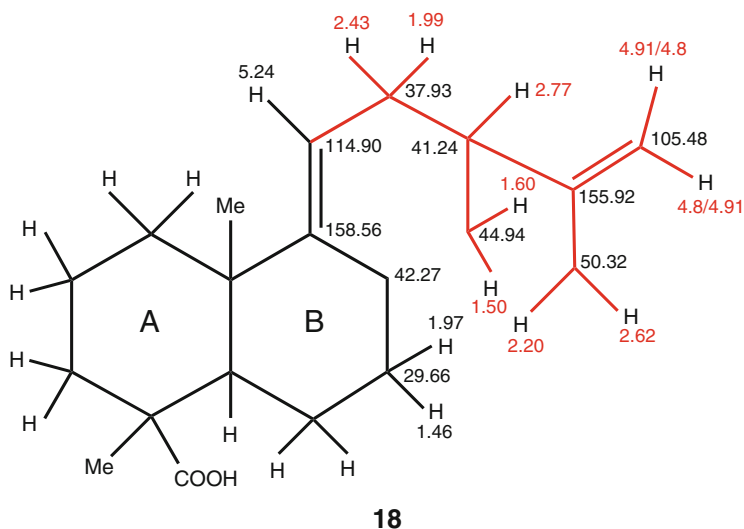
Continuing elucidation of the unknown compound, fragment **11** contains the same proton (1.67 ppm) for which the attached methine carbon (46.56 ppm) exhibited HMBC connectivities to both methyl groups. This fragment can then be added to the developing structure by means of HMBC correlations to produce **16**. HMBC data reveal that the proton at 1.68 ppm shows correlations to the carbons at 44.69 and 38.79 ppm and that at 2.47 ppm displays connectivities to the carbons at 46.56 and 44.69 ppm.



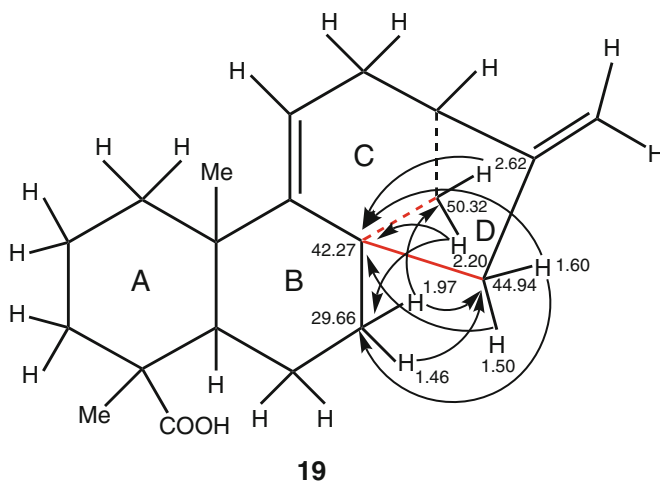
The next step in this structural determination is closure of an apparent ring. Protons in **16** at 2.47, 1.97, and 1.46 ppm exhibit HMBC correlations to a carbon at 42.27 ppm and suggest that it be placed between the carbons at 158.56 and 29.66 ppm to complete the B-ring (**17**). Supporting evidence comes from the alkenic, methine proton at 5.24 ppm, which is attached to the carbon at 114.90 ppm and shows HMBC connectivities to the previously identified carbons at 158.56 ppm (two-bond) and 38.79 ppm (three-bond). It also exhibits a strong HMBC correlation to the above carbon at 42.27 ppm, which is consistent with this quaternary carbon being located in an (*E*)-alkenic position between the carbons at 158.56 and 29.66 ppm, thus supporting structure **17**.



The alkenic methine proton at 5.24 ppm and its attached carbon at 114.90 ppm are also part of the fragment shown in **12**. Adding this large piece to structure **17** produces **18**, which contains five of the seven units of unsaturation.



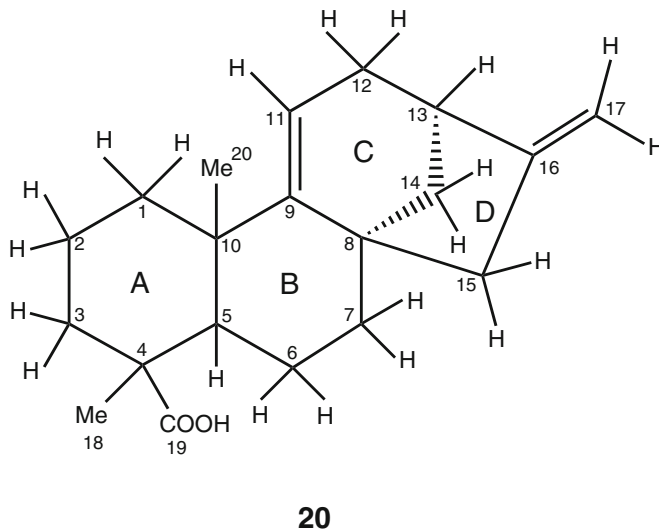
Closure of an apparent ring again requires additional HMBC connectivities. In particular, (i) the B-ring methylene group (protons at 1.97 and 1.46 ppm, carbon at 29.66 ppm) and (ii) the two newly added methylene groups (1.60 and 1.50 ppm/44.94 ppm and 2.62 and 2.20 ppm/50.32 ppm) appear well positioned to complete the unknown molecular structure. First, the protons at 2.62, 2.20, 1.60, and 1.50 ppm show HMBC connectivities to the quaternary carbon at 42.27 ppm, thus closing the C- and D-rings (**19**).



Second, the proton at 1.97 ppm displays HMBC correlations to both methylene carbons at 44.94 and 50.32 ppm, and the proton at 1.46 ppm exhibits an HMBC

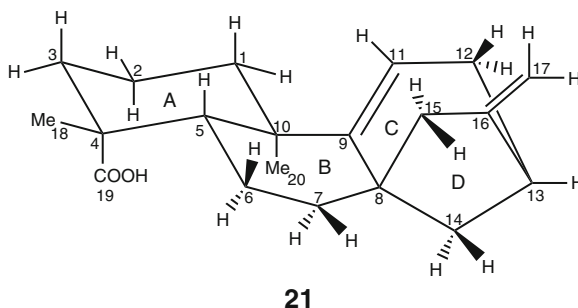
connectivity to the carbon at 44.94 ppm. In addition, the protons at 2.20 and 1.60 ppm display complementary HMBC correlations to the carbon at 29.66 ppm.

At this point all of the carbons, hydrogens, and oxygens are accounted for. The complete, numbered, structure **20** of the unknown compound satisfies the required seven units of unsaturation.



7.6 Determination of Overall Stereochemistry and Proton Chemical Shift Assignments

With completion of the two-dimensional structural elucidation of the unknown compound, questions arise concerning its three-dimensional shape, *i.e.* the relative orientations of substituents (*e.g.* whether methyl-18 at C-4 is *axial* or *equatorial*) at various carbons. For a molecule of MW = 300, a NOESY experiment can provide a wealth of such stereochemical information. The NOESY spectrum of the “unknown,” kauradienoic acid (**3**), is shown in Fig. 7. The data from these and ROESY contour plots are summarized in Table 5 and illustrated, in part, in a numbered, stereochemical drawing (**21**).



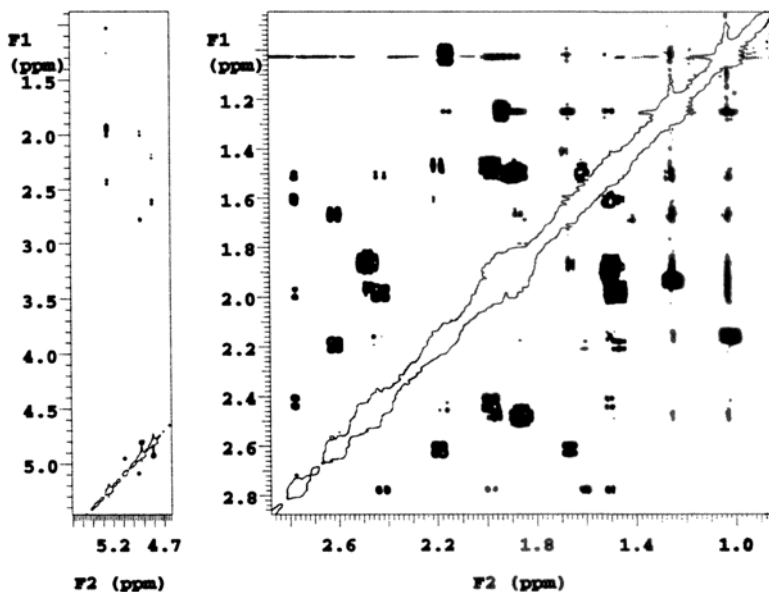


Fig. 7 NOESY spectrum for kauradienoic acid (3) with a 0.5-s mixing time. The expansions are similar to those in Fig. 2

Table 5 NOE and ROE data for kauradienoic acid (3)

δ_H /ppm	Dipolar-coupled partners
5.24	2.43, 1.99, 1.93, 1.24, 1.02
4.91 ^a	4.80, 2.77, 1.99
4.80 ^b	4.91, 2.62, 2.20
2.77	4.91, 2.43, 1.99, 1.60, 1.50
2.62	4.80, 2.20, 1.68, 1.67
2.47	1.97, 1.68, 1.46, 1.24, 1.02
2.43	5.24, 2.77, 1.99, 1.50
2.20	4.80, 2.62, 1.60, 1.46
2.16	1.88, 1.50, 1.24, 1.02
1.99	5.24, 4.91, 2.77, 2.43
1.97	2.47, 1.50, 1.46, 1.02
1.93	5.24, 1.50, 1.24, 1.02
1.88	2.16, 1.50, 1.02
1.68	2.62, 2.47, 1.46
1.67	2.62, 1.24, 1.02
1.60	2.77, 2.20, 1.50
1.50 ^c	2.77, 2.43, 2.16, 1.97, 1.93, 1.88, 1.60,
1.50 ^c	1.24, 1.02
1.46	2.47, 2.20, 1.97, 1.68
1.24 ^d	5.24, 2.47, 2.16, 1.93, 1.67, 1.50, 1.02
1.24 ^d (CH ₃)	
1.02 ^e	5.24, 2.47, 2.16, 1.97, 1.93, 1.88,
1.02 ^e (CH ₃)	1.67, 1.50, 1.24

^a*cis* to C-13

^b*cis* to C-15

^{c,d,e}Correlations cannot be differentiated for the pairs of hydrogens at 1.50, 1.24, and 1.02 ppm because of their identical chemical shifts

Since the 300 molecular weight of kauradienoic acid (**3**) is well outside of the zero-crossing limit, the NOESY and ROESY spectra were expected to be virtually identical, an expectation that was borne out, and subsequent references will be to "NOESY" data only. A combination of NOESY and HMBC spectra can permit the determination of substituent stereochemistry, especially for 6-membered ring systems that exist in chair conformations, in the following way: actual or quasi-*equatorial* protons exhibit strong HMBC correlations by virtue of their $\sim 180^\circ$ dihedral angles while actual or quasi-*axial* protons display strong NOEs with other *axial* protons and methyl groups because of their proximity (note that NOEs vary with the sixth power of the internuclear distance).

The following overall structural relationships were thus deduced for kauradienoic acid. Strong NOEs among (i) H-1 β ,*ax.*; H-3 β ,*ax.*; and H-5 β ,*ax.* and (ii) between H-2 α ,*ax.* and methyl-20*ax.* and strong HMBC correlations between (i) H-1 α ,*eq.* and C-3 and C-5; H-2 β ,*eq.* and C-4 and C-10; H-3 α ,*eq.* and C-1 and C-5, (ii) H-1 β ,*ax.* and C-20, and (iii) H-3 β ,*ax.* and C-19 demonstrate that the A-ring exists in a chair conformation in which H-1 β , H-3 β , and H-5 β are on the "top face" of the molecule while H-2 α and methyl-20 are on the "bottom face."

X-ray data show that the B-ring occurs as a slightly distorted boat conformer, in which methyl-20 and H-7 α are at the "flagpole" positions. However, this conformation could have been reasonably inferred from the strong NOEs that are observed between these flagpole groups and the strong HMBC connectivities (due to $\sim 180^\circ$ dihedral angles) that are seen between H-6 α and C-8 and between H-7 β and C-5. *Dreiding* models suggest that the B-ring might also exist partially as a quasi-half chair conformer, in which C-6 would be well below the C-5–C-10–C-9–C-8 plane and C-7 slightly below this plane. However, this conformation places methyl-20 relatively distant from H-7 α (1.97 ppm) and much closer to H-6 α (2.47 ppm). The relatively weak NOE observed between methyl-20 and H-6 α and strong NOE seen for methyl-20 and H-7 α indicate that the distorted half-chair conformer is unimportant in solution.

Strong HMBC correlations between H-11 and C-8 and C-13, and H-13 and C-11 are consistent with the C-ring occurring in a rigid conformation in which carbons 8, 9, 11, 12, and 13 are approximately coplanar and C-14 well below the plane. Additional strong HMBC connectivities between H-14 β and C-9 and C-12 and between H-12 β and C-14 support this inference.

The remaining 5-membered D-ring is situated approximately orthogonally to the A,B,C-ring system, and descriptions of the 15-protons as being " α " or " β " must, therefore, be clarified (in Table 6, *vide infra*, they are described as pointing "toward" or "away from" the viewer, respectively). Strong HMBC correlations between H-14 α and C-15 and C-16; H-15 α and C-9; H-15 β and C-14; and H-12 α and C-16 support this conclusion. One sees then that H-14 β is quasi-*equatorial* with respect to the C-ring while H-14 α is quasi-*equatorial* with respect to the D-ring.

Proton assignments were either made or confirmed on the basis of NOE data. Strong NOESY correlations between the protons at 1.97 ppm (at C-7) and 1.50 ppm

(at C-14) indicate that both are similarly oriented. Since the 7-*geminal* partner at 1.46 ppm has been shown, *via* the HMBC experiments, to be β , the H-7 at 1.97 ppm must, therefore, be α . Thus, H-14 at 1.50 ppm is confirmed to be α by virtue of its strong NOE interaction with H-7 α . NOESY connectivities were also used to assign the alkenic methylene protons at C-17. H-17A (4.91 ppm) shows a strong NOE with H-13 (2.77 ppm) and is thus *cis* to it while H-17B (4.80 ppm) displays equally strong NOEs to the *cis* 15-protons at 2.62 and 2.20 ppm.

Finally, two additional pieces of evidence support the previous assignment of methyl-20 at 1.02, not 1.24 ppm, and protons 1 β and 3 β at 1.24 and 1.02 ppm, respectively, both of which can be seen in **22**: (i) strong NOEs (shown as arrows) that can exist only between methyl-20 and H-2 α . (1.88 ppm) and H-7 α (1.97 ppm) and (ii) HMBC traces through carbons 9 (158.56 ppm) and 19 (COOH, 182.74 ppm) show particular emphasis on the fine structure of H-1 α . (1.24 ppm) and H-3 α . (1.02 ppm) (Fig. 8).

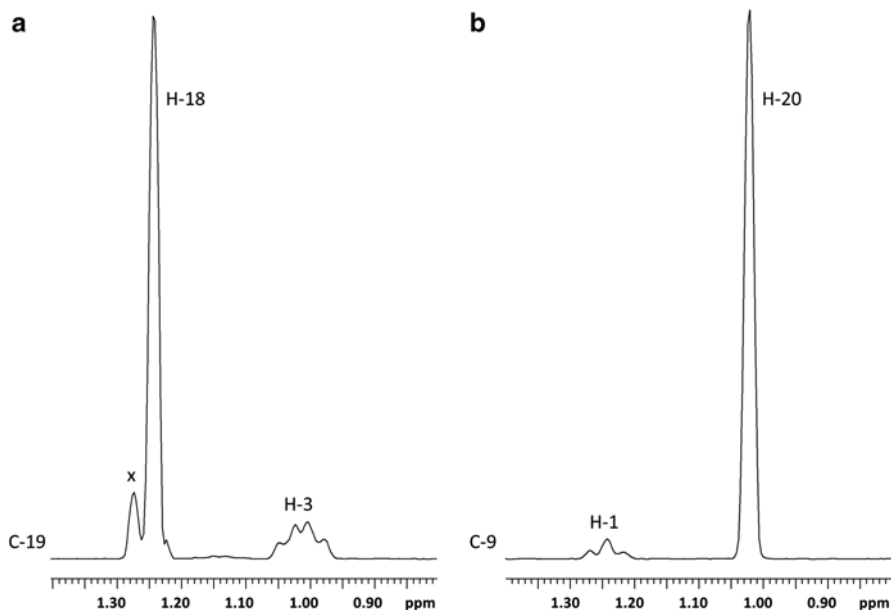
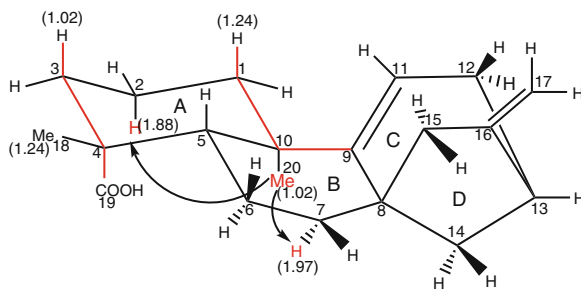


Fig. 8 Cross-sections through C-9 and C-19 from an HMBC spectrum of kauradienoic acid (**3**). The observation of multiplet structures for H-1 and H-3 *axial* protons allows one to clearly distinguish these two protons from the H-18 and H-20 methyl singlets, which overlap with these protons in the ^1H spectrum. The peak marked "x" is a minor methyl impurity peak. H-3 appears as an apparent quartet due to a *geminal* H–H coupling to H-3 *eq.*, an *anti* H–H coupling to H-2 *ax.*, and an *anti* C–H coupling to C-19. H-1 appears as an apparent triplet due to a *geminal*-coupling to H-1 *eq.* and an *anti*-coupling to H-2 *ax.* In both cases, the resolution is insufficient to observe *gauche* H–H couplings and the *gauche* C–H coupling of H-1 to C-9



22

H-3 $_{ax}$ appears as a triple doublet (an apparent 1:3:3:1 “quartet”) by virtue of the magnitude of its similar *geminal*-coupling to H-3 $_{eq}$, *axial*-coupling to H-2 $_{ax}$, and *axial* C-H coupling to 19-COOH, the third pathway shown in red in **22**. However, H-1 $_{ax}$ is seen as a doublet of doublets (an apparent 1:2:1 “triplet”) due to its similar *geminal*-coupling to H-1 $_{eq}$ and *axial*-coupling to H-2 $_{ax}$. Its third coupling is *equatorial* to C-9 (also shown in red) and too small to cause observable splitting. With the interpretation of these final HMBC and NOESY correlations, the two- and three-dimensional structural elucidation of kauradienoic acid (**3**) is complete, and final descriptions of the skeletal protons are given in Table 6.

Table 6 ^1H and ^{13}C NMR data for kauradienoic acid (**3**)

Position	δ_c /ppm	Attached hydrogens
1	40.75	1.93 $\alpha, eq.$ 1.24 $\beta, ax.$
2	20.16	1.50 $\beta, eq.$ 1.88 $\alpha, ax.$
3	38.31	2.16 $\alpha, eq.$ 1.02 $\beta, ax.$
4	44.69	–
5	46.56	1.67 $\beta, ax.$
6	18.48	1.68 β 2.47 α
7	29.66	1.46 β 1.97 α
8	42.27	–
9	158.56	–
10	38.79	–
11	114.90	5.24
12	37.93	1.99 β 2.43 α
13	41.24	2.77
14	44.94	1.60 β 1.50 α
15	50.32	2.62 ^a 2.20 ^b
16	155.92	–
17	105.48	4.91 ^c 4.80 ^d
18	28.23	1.24 $\beta, eq.$ (CH ₃)
19	182.74	–
20	23.62	1.02 $\alpha, ax.$ (CH ₃)

^aPointing away from the viewer

^bPointing toward the viewer

^c*cis* to C-13

^d*cis* to C-15

8 Computer-Assisted Structure Elucidation

By the late 1950s, chemists realized that NMR spectroscopy was a powerful tool for identifying the structure of organic compounds. The first computer-assisted structure elucidation (CASE) programs, for small organic compounds, were developed in the late 1960s: DENDRAL (79), CHEMICS (80), CASE (81), and StRec (82). However, early elucidation attempts were severely limited by a number of factors, not the least of which was the relatively primitive nature of computers at that time. In addition, chemists were essentially limited to ^1H NMR spectroscopy, as practical one-dimensional ^{13}C and two-dimensional NMR experiments were years away.

The situation changed dramatically with the advent of powerful personal computers, ^{13}C NMR spectroscopy, and a variety of 2D NMR experiments to establish direct and long-range homo- and heteronuclear connectivities. A second generation of considerably more powerful structural elucidation programs appeared several decades later and included Structure Elucidator (56), SESAMI (83), LSD (84), CISOC-SES (85), LUCY (86), and COCON (87).

One of the greatest difficulties in elucidating the structure of an unknown compound arises when spectroscopic data are consistent and *appear* to lead to one structure. This problem is especially acute if a chemist has isolated, or been given, similar compounds in the past, and the current unknown seems to be another analog in this series. The obvious advantage of CASE programs is that they do not suffer a similar bias. Some of their proposed structures may be highly implausible, but they can also suggest classes of structures that the chemist has not even considered.

Section 7 has illustrated a typical workflow for the elucidation of an unknown structure in the absence of a computer-assisted elucidation. Advanced Chemistry Development, Ltd. (ACD/Labs) markets a program, “ACD/Structure Elucidator” (56); by comparison, their automated determination of structures occurs in the following general manner.

(1) Spectral data requirements:

- (a) ^1H NMR data: useful for peak integral information.
- (b) $^1\text{H}/^{13}\text{C}$ HSQC, $^1\text{H}/^1\text{H}$ COSY, and $^1\text{H}/^{13}\text{C}$ HMBC.
- (c) $^1\text{H}/^1\text{H}$ TOCSY: useful when COSY spectra indicate the presence of complex spin systems.
- (d) ^{13}C NMR data: very helpful for identifying quaternary carbons, if sample quantities permit its acquisition in a reasonable amount of time and can be critical for quaternary carbons that do not exhibit HMBC connectivities.
- (e) High-resolution MS to provide a molecular formula.
- (f) IR and UV/vis spectral data to furnish information on functional groups.

(2) NMR data are submitted in one of the following two formats:

- (a) As raw one- and two-dimensional spectral data (FIDs together with their processing parameters) in which 1D and 2D NMR cross-peaks are selected by the software. This is preferable because data entry is both rapid and direct.
- (b) As text (TXT) files in which the chemist/spectroscopist has analyzed the NMR spectra to establish various homo- and heteronuclear connectivities.

The latter approach is less favored because it can involve laborious data entry and introduces the serious possibility of transcription errors.

- (3) Numerous conditions can be applied to NMR data processing, but as a general rule, the following are applied:
 - (a) HMBC correlations are assigned as follows:
 - (i) Strong: ${}^2J(\text{CH}) - {}^3J(\text{CH})$
 - (ii) Weak: ${}^2J(\text{CH}) - {}^4J(\text{CH})$
 - (b) Direct heteroatom-to-heteroatom connectivities are disallowed
 - (c) Triple bonds within 3-, 4-, and 5-membered rings are disallowed
- (4) A molecular-connectivity map is then generated, which shows all of the NMR data in one place and is correlated to the molecular formula. The NMR data include heteronuclear (HMBC) and homonuclear (COSY and, possibly, TOCSY) connectivities. While interesting, these maps are generally too complicated to be used alone to produce possible structures.
- (5) Potential structures, numbering in the tens or hundreds, are generated and ${}^{13}\text{C}$ chemical shifts calculated for each candidate structure. Differences between predicted and experimental data are reported as a “fast-deviation” statistic, $d_{\text{F}}({}^{13}\text{C})$, and possible structures initially ranked in order of increasing $d_{\text{F}}({}^{13}\text{C})$.
- (6) More accurate ${}^{13}\text{C}$ chemical shifts calculations are next performed on the smaller of either (i) all structures with $d_{\text{F}}({}^{13}\text{C}) \leq 4$ ppm/carbon or (ii) the first 50 ranked structures. Differences between the more accurately predicted and experimental data are reported as an “accurate-deviation” statistic, using “neutral net” [$d_{\text{N}}({}^{13}\text{C})$] values and/or “HOSE” (Hierarchically Ordered Spherical Description of Environment) [$d_{\text{A}}({}^{13}\text{C})$]. The two methods give somewhat similar results, but the HOSE method yields better results when compounds, which are similar to the unknown, are contained in the ACD spectral library. Potential structures are re-ranked in order of increasing $d_{\text{A/N}}({}^{13}\text{C})$ numbers. Structures having $d_{\text{A/N}}({}^{13}\text{C})$ values > 4 ppm/carbon are generally discarded. ACD/Labs’ experience is that the correct structure is usually identified at this point (88).
- (7) In situations where the smallest calculated $d_{\text{A/N}}({}^{13}\text{C})$ values are very close, a best structure may be arrived at by calculating $d_{\text{A/N}}({}^1\text{H})$, and, if good MS data are available, $d(\text{MS})$ can be calculated as well. In the following sections, three test compounds were submitted to ACD/Labs to determine how the performance of their ACD/Structure Elucidator program would compare to that of an experienced NMR spectroscopist.

8.1 Guyanin

The structure of guyanin (**4**) was determined in 1986 (32, 89) and represents one of the first and most unusual structures determined solely by NMR methods. The structure was so unprecedented that one of the senior co-authors felt that confirming

X-ray data should be obtained. The results of X-ray analysis finally arrived, just prior to submission of the manuscript, and completely supported the NMR-determined structure.

Critical 2- and 3-bond H-C connectivities were established by the XCORFE experiment, which preceded development of the HMBC experiment. Since the original NMR data for guyanin (**4**) are no longer available, the following data tables were reconstructed from data in the original manuscripts and submitted to ACD/Labs as TXT files: HETCOR (direct H-C chemical-shift correlation (Table 7), COSY (Table 8), and XCORFE (longer-range H-C chemical-shift correlation, Table 9).

Table 7 HETCOR^a data for guyanin (**4**)

δ_c /ppm	Attached hydrogens (with coupling constants)	
208.41	–	
176.95	–	
175.28	–	
173.67	–	
173.24	–	
164.25	–	
142.86	7.44 t (1.7)	
140.25	7.56 t (1.7)	
128.89	5.87 q (1.3)	
120.85	–	
117.45	5.90 s	
108.51	6.38 dd (1.7, 0.8)	
79.95	5.94 dd (1.7, 0.8)	
68.95	–	
61.45	–	
51.98	3.68 s	
51.25	3.60 s	
46.57	–	
46.30	–	
40.91	3.26 d (10)	
34.39	1.71 td (13, 13, 6)	1.56 dd (13, 6)
34.33	2.55 dd (17, 10)	2.19 d (17)
28.31	2.29 ddd (15, 13, 6)	2.16 dd (15, 6)
27.80	0.98 s	
23.65	1.27 s	
20.67	1.43 s	
18.99	1.16 s	
16.37	2.04 d (1.3)	

^aA carbon-detected, direct heteronuclear chemical shift correlation experiment

Table 8 COSY data for guyanin (**4**)

δ_H/ppm	Spin-coupled partners
7.56	7.44, 5.94
7.44	7.56, 6.38
6.38	7.44, 5.94
5.94	7.56, 6.38
5.90	(none)
5.87	2.04
3.68 (CH ₃)	(none)
3.60 (CH ₃)	(none)
3.26	2.55
2.55	3.26, 2.19
2.29	2.16, 1.71, 1.56
2.19	2.55
2.16	2.29, 1.71
2.04 (CH ₃)	5.87
1.71	2.29, 2.16, 1.56
1.56	2.29, 1.71
1.43 (CH ₃)	(none)
1.27 (CH ₃)	(none)
1.16 (CH ₃)	(none)
0.98 (CH ₃)	(none)

The Structure Elucidator program generated a molecular connectivity diagram (Fig. 9) and a single structure (**23**). As it turned out, the COSY data were not needed. Only one structure, the correct one, was produced with a generation time of 1 s. Table 10 contains ¹³C and ¹H chemical shift data that are sorted by position number.

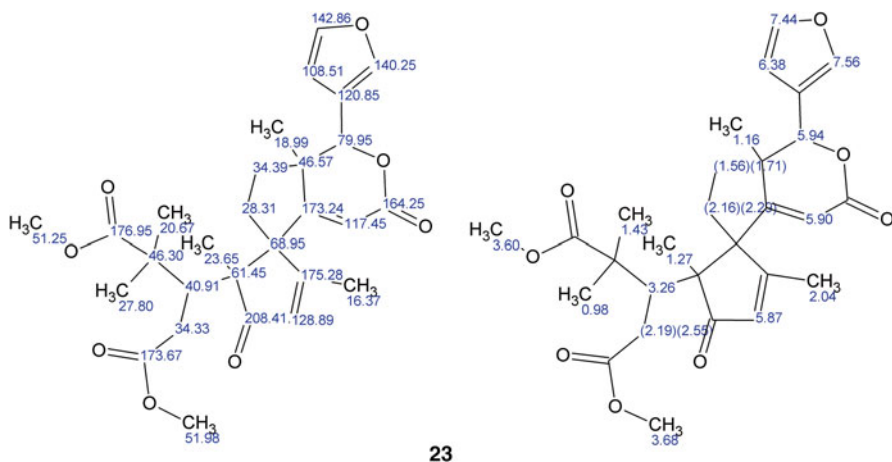


Table 9 XCORFE^a data for guyanin (4)

δ_c /ppm	Long-range correlated hydrogens
208.41	5.87, 3.26, 2.04, 1.27
176.95	3.60, 3.26, 1.43, 0.98
175.28	5.87, 2.29, 2.04
173.67	3.68, 3.26, 2.55, 2.19
173.24	5.94, 5.90, 2.16, 1.56, 1.16
164.25	5.94, 5.90
142.86	7.56, 6.38
140.25	7.44, 6.38, 5.94
128.89	2.04
120.85	7.56, 7.44, 6.38, 5.94
117.45	(none)
108.51	7.56, 7.44, 5.94
79.95	1.71, 1.16
68.95	5.90, 5.87, 3.26, 2.29, 2.16, 2.04, 1.56, 1.27
61.45	5.87, 3.26, 2.29, 2.19, 2.16, 1.27
51.98	(none)
51.25	(none)
46.57	5.94, 5.90, 2.16, 1.71, 1.56, 1.16
46.30	3.26, 2.55, 2.19, 1.43, 0.98
40.91	2.55, 2.19, 1.43, 1.27, 0.98
34.39	2.29, 2.16, 1.16
34.33	3.26
28.31	1.71, 1.56
27.80	3.26, 1.43
23.65	3.26
20.67	3.26, 0.98
18.99	5.94, 1.71, 1.56
16.37	5.87

^aA carbon-detected, indirect heteronuclear chemical shift correlation experiment

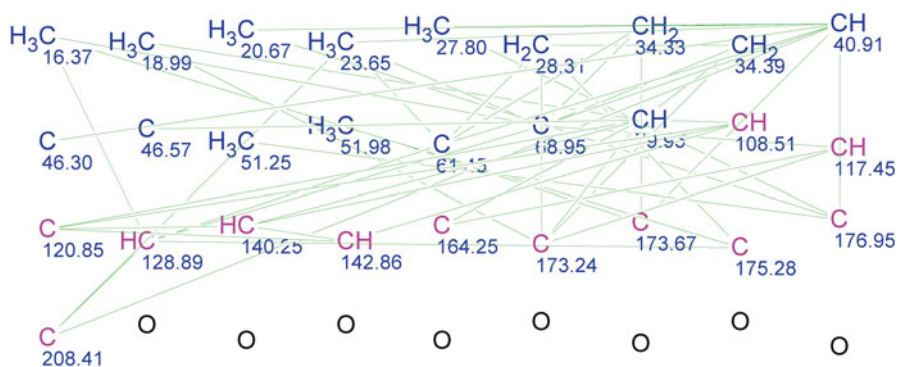


Fig. 9 An ACD molecular connectivity diagram for guyanin (4) showing the various carbon–carbon connections

Table 10 ^1H and ^{13}C NMR data for guyanin (**4**)

Position	$\delta_{\text{C}}/\text{ppm}$	Attached hydrogens	
1	176.95	–	
2	46.30	–	
3	40.91	3.26 d (10)	
4	34.33	2.55 dd (17, 10)	2.19 d (17)
5	173.67	–	
6	61.45	–	
7	208.41	–	
8	128.89	5.87 q (1.3)	
9	175.28	–	
10	68.95	–	
11	28.31	2.29 ddd (15,13,6)	2.16 dd (15, 6)
12	34.39	1.71 td (13,13,6)	1.56 dd (13, 6)
13	46.57	–	
14	173.24	–	
15	117.45	5.90 s	
16	164.25	–	
17	79.95	5.94 dd (1.7, 0.8)	
18	120.85	–	
19	108.51	6.38 dd (1.7, 0.8)	
20	142.86	7.44 t (1.7)	
21	140.25	7.56 t (1.7)	
22	18.99	1.16 s	
23	16.37	2.04 d (1.3)	
24	23.65	1.27 s	
25	27.80	0.98 s	
26	20.67	1.43 s	
27	51.25	3.60 s	
28	51.98	3.68 s	

8.2 T-2 Toxin

The sesquiterpene T-2 toxin is a member of the trichothecene family of mycotoxins. Its structure (**5**) was determined in 1968 (90). Proton and ^{13}C NMR data were collected at 400 and 100 MHz, as a teaching aid, to illustrate the structural elucidation of medium-sized organic molecules and to check the assignments, with regard to relative orientations, of various protons within the molecule (40).

In this case, raw one- and two-dimensional spectral data (FIDs and their processing parameters) were submitted to ACD/Labs for analysis. Summaries of these data are given in the following tables: HSQC (Table 11), COSY (Table 12), and HMBC (Table 13).

The Structure Elucidator program generated a molecular connectivity diagram (Fig. 10) and four possible structures (**24**). Of these four, only the two structures with epoxy groups (**25** and **26**) passed filtering, *viz.* their $d_{\text{A}}(^{13}\text{C})$ values, used here because they are more discriminating than the corresponding $d_{\text{N}}(^{13}\text{C})$ values, are less

Table 11 HSQC data for T-2 toxin (**5**)

δ_c /ppm	Attached hydrogens (with coupling constants)	
173.00	–	
172.80	–	
170.39	–	
136.59	–	
123.77	5.81 d (5.0)	
84.68	5.31 d (2.4)	
78.81	3.70 d (5.0)	
78.48	4.16 ddd (5.0, 2.4, 2.4 ^a)	
68.13	5.29 d (5.5)	
67.43	4.35 d (5.0)	
64.69	4.29 d (12.6)	4.06 d (12.6)
64.60	–	
48.68	–	
47.30	3.07 d (4.0)	2.80 d (4.0)
43.67	2.16 d (6.3) [2H]	
43.25	–	
27.87	2.41 dd (14.0, 5.5)	1.91 d (14.0)
25.88	2.10 m [1H]	
22.55	0.97 d (6.4) [CH ₃]	
22.48	0.96 d (6.4) [CH ₃]	
21.15	2.04 s [CH ₃]	
21.10	2.15 s [CH ₃]	
20.45	1.75 s [CH ₃]	
6.95	0.81 s [CH ₃]	
–	3.19 d (2.4 ^a) [OH]	

^aSpin coupling observed between these signals in a fresh sample

Table 12 COSY data for T-2 toxin (**5**)

δ_H /ppm	Spin-coupled partners
5.81	5.29, 4.35, 1.75
5.31	4.16
5.29	5.81, 2.41, 1.75
4.35	5.81, 4.06, 1.91, 1.75
4.29	4.06, 2.41
4.16	5.31, 3.70, 3.19 ^a
4.06	4.35, 4.29
3.70	4.16
3.19 [OH]	4.16 ^a
3.07	2.80
2.80	3.07
2.41	5.29, 4.29, 1.91
2.16 [2H]	2.10
2.15 [CH ₃]	(none)
2.10	2.16, 0.97, 0.96
2.04 [CH ₃]	(none)
1.91	4.35, 2.41
1.75 [CH ₃]	5.81, 5.29, 4.35
0.97 [CH ₃]	2.10
0.96 [CH ₃]	2.10
0.81 [CH ₃]	(none)

^aSpin coupling is observed between these protons in a fresh sample

Table 13 HMBC data for T-2 toxin (5)

δ_c /ppm	Longer-range correlated hydrogens
173.00	5.31, 2.04
172.80	5.29, 2.16
170.39	4.29, 4.06, 2.15
136.59	5.29, 4.35, 1.91, 1.75
123.77	5.29, 4.35, 1.75
84.68	4.16, 3.70, 0.81
78.81 ^a	5.31, 3.07, 2.80
78.48 ^a	5.31, 3.70
68.13	5.81, 4.35, 2.41, 1.91, 1.75
67.43	5.81, 4.29, 4.06, 3.70, 1.91, 1.75
64.69	5.31, 4.35, 2.41, 1.91
64.60	5.31, 4.29, 4.06, 3.70, 3.07, 2.80, 0.81
48.68	5.31, 4.35, 4.29, 4.06, 3.70, 3.07, 2.80, 2.41, 1.91, 0.81
47.30	3.70
43.67	2.10, 0.97, 0.96
43.25	5.81, 5.31, 5.29, 4.35, 4.29, 4.06, 2.41, 1.91, 0.81
27.87	5.29, 4.35, 4.29, 4.06
25.88	2.16, 0.97, 0.96
22.55	2.16, 2.10, 0.96
22.48	2.16, 2.10, 0.97
21.15	(none)
21.10	(none)
20.45	5.81, 5.29, 4.35
6.95	3.70

^aHMBC connectivity is observed to the OH signal ($\delta_{3.19}$) in a fresh sample

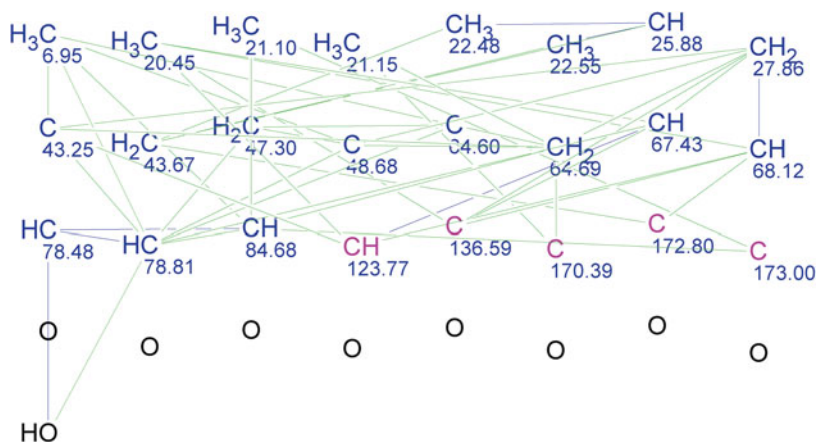
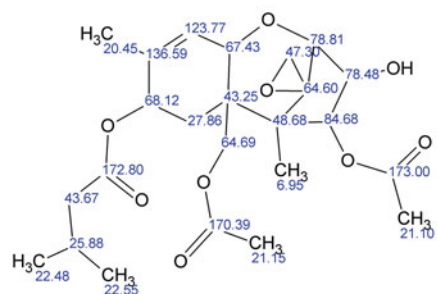


Fig. 10 An ACD molecular connectivity diagram for T-2 toxin (5) showing the various carbon-carbon connections

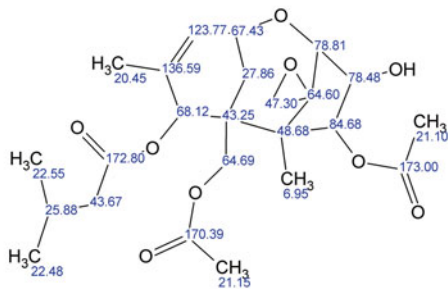
than 4 ppm. These structures differ in the placement of the methylene carbon at 27.86 ppm. However, **25** had a considerably better $d_A(^{13}\text{C})$ value than **26** (0.816 vs. 2.187 ppm) and proved to be correct.

1

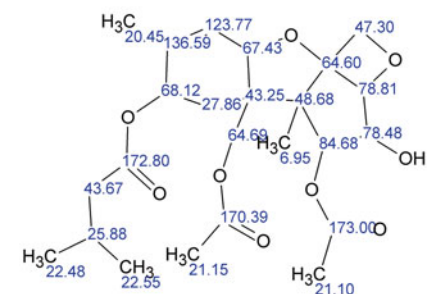
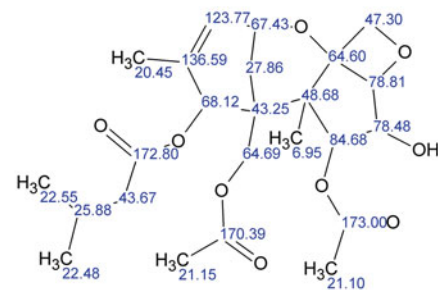
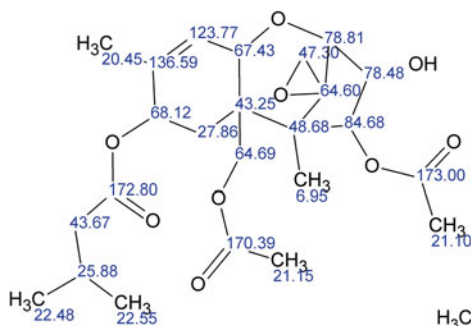
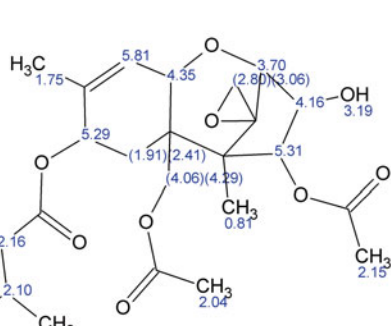
 $d_A(^{13}\text{C})$: 0.816

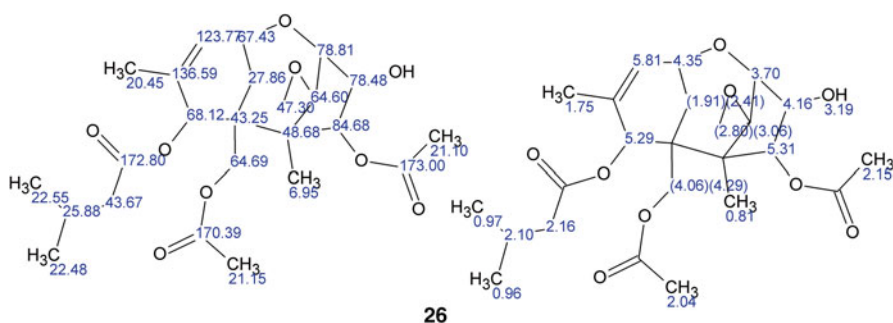
3

2

 $d_A(^{13}\text{C})$: 2.187

4

 $d_A(^{13}\text{C})$: 4.095 $d_A(^{13}\text{C})$: 4.459**24****25**



It might be surprising that the incorrect structure, which involves the shift of a methylene carbon, would score as well as it did. Part of the reason is that critical HMBC correlations are observed in both structures, being two-bond in one and three-bond in the other, which cannot be distinguished. The presence of a key COSY correlation between the protons at 2.41 and 5.29 ppm permits identification of the correct structure. Table 14 contains ^{13}C and ^1H chemical shift data that are sorted by position number.

Table 14 ^1H and ^{13}C data for T-2 toxin (**5**)

Position	δ_{C} /ppm	Attached hydrogens
2 ^a	78.81	3.70 d (5.0)
3	78.48	4.16 ddd (5.0, 2.4, 2.4 ^b)
4	84.68	5.31 d (2.4)
5	48.68	—
6	43.25	—
7	27.87	2.41 dd (14.0, 5.5) 1.91 d (14.0)
8	68.13	5.29 d (5.5)
9	136.59	—
10	123.77	5.81 d (5.0)
11	67.43	4.35 d (5.0)
12	64.60	—
13	47.30	3.07 d (4.0) 2.80 d (4.0)
14	6.95	0.81 s
15	64.69	4.29 d (12.6) 4.06 d (12.6)
16	20.45	1.75 s
1'	172.80	—
2'	43.67	2.16 d (6.3) [2H]
3'	25.88	2.10 m
4'	22.55	0.97 d (6.4)
5'	22.48	0.96 d (6.4)
4-OAcCH ₃	21.15	2.04
15-OAcCH ₃	21.10	2.15
4-OAc	173.00	—
15-OAc	170.39	—
		3.19 d (2.4 ^b) [OH]

^aPosition 1 is assigned to the ether oxygen between carbons 2 and 11

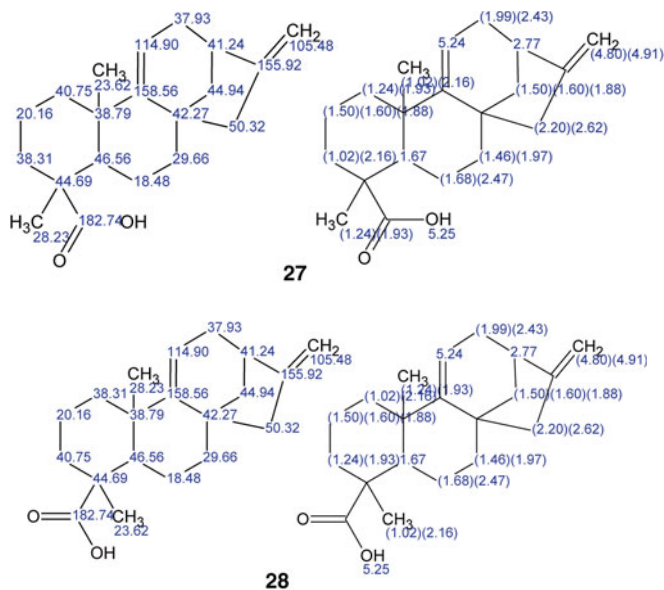
^bSpin coupling is observed between these protons in a fresh sample

8.3 Kauradienoic Acid

As mentioned in Sect. 7, kauradienoic acid is a diterpene, for which the structure **3** was determined in 1971 (78) and its ^1H and ^{13}C NMR spectra completely assigned in 1984 (31). This compound was used earlier in the section to illustrate how a general structural elucidation could be achieved through the systematic analysis of ^1H , ^{13}C , HSQC, COSY, TOCSY, and HMBC spectra. The process is somewhat laborious, even for a relatively small-molecular weight compound, due to the occurrence of many close-lying proton NMR signals.

Raw one- and two-dimensional spectral data (FIDs and their processing parameters) were again submitted to ACD/Labs for analysis. Summaries of these data are included in the following tables: HSQC (Table 1), COSY (Table 2), TOCSY (Table 3), and HMBC (Table 4), which are given in Sect. 7.

The ACD/Structure Elucidator program generated a molecular connectivity diagram (Fig. 11) and two structures, **27** and **28**. Inspection of the two structures reveals the same problem that was encountered in Sect. 7, *viz.* that ^1H and ^{13}C assignments for two methyl groups and their two adjacent methylene groups are interchanged. In this case, use of $d_{\text{N}}(^{13}\text{C})$ values gave slightly better discrimination. However, it was not surprising that very similar numbers, 1.17 and 1.26 ppm, were calculated for identical structures with three differing assignments. Again, the correct structure was the lower one. Assignment of the proton chemical shifts was made especially difficult due to accidental chemical shift equivalence of three sets of protons: H-3 β and methyl-20 (1.02 ppm), H-1 β and methyl-18 (1.24 ppm), and H-2 β and H-14 α (1.50 ppm) (Table 6). As a result, structures **27** and **28** both exhibit composite assignments in which the assignments of ^1H chemical shifts for methyl (1.02 and 1.24 ppm) and certain methylene protons (1.50, 1.60, 1.88, 1.93, and 2.16 ppm) can be interchanged.



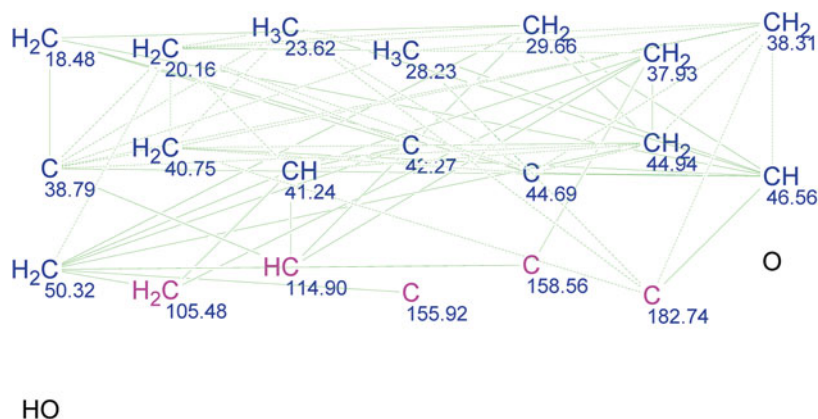


Fig. 11 An ACD molecular connectivity diagram for kauradienoic acid (**3**) showing the various carbon-carbon connections

However, as was seen at the conclusion of Sect. 7.6, several lines of reasoning based on the analysis NOESY connectivities and HMBC traces demonstrated unequivocally that only the ^1H and ^{13}C assignments shown in **13** (in Sect. 7.5) and **27** (in this section) could be correct. ^{13}C and ^1H chemical shift data for kauradienoic acid (**3**), which are sorted by position number, are presented in Table 6 in Sect. 7.6.

The examples, illustrated in Sects. 8.1–8.3, demonstrate that the ACD/Structure Elucidator program is an important addition to the arsenal of chemists engaged in the determination of structures of organic compounds. Moreover, it is best when used in conjunction with a knowledgeable NMR spectrometer operator, who can distinguish between different structural possibilities by closer examination of the NMR spectra.

9 The Effect of Dynamic Processes on the Appearance of NMR Spectra of Natural Products and Other Organic Compounds

There are a number of exchange processes (conformational interchange, tautomerism, epimerization, *etc.*), which can affect the appearance of an NMR spectrum. In considering how these processes can alter the appearance of an NMR spectrum, it is important to recognize that what matters is the frequency difference between a pair of peaks interchanged by the exchange process, compared to the frequency of exchange. For this reason, the present authors dislike the commonly used term “NMR time scale” since this implies a single time scale for a spectrum measured on a particular spectrometer. While the peak separations for individual pairs of peaks are directly proportional to the spectrometer operating frequency, different pairs of exchanging peaks for the same compound will generally have different peak

separations, *i.e.* there are several different “time scales” for a single compound at a single acquisition frequency. A further objection to this term is seen when considering ^{13}C and ^1H spectra run on the same spectrometer. While the acquisition frequency for the ^{13}C spectrum is almost exactly $\frac{1}{4}$ of that for ^1H , the ^{13}C chemical shift range (in ppm) is typically *ca.* 20 times that for ^1H . Consequently, peak separations (in frequency units) for pairs of exchanging ^{13}C peaks are typically significantly larger than the peak separations for corresponding ^1H pairs. Thus, rather than having a longer “NMR time scale” (as would be implied by the lower frequency), ^{13}C spectra actually usually have shorter “time scales,” *i.e.* faster exchange rates are needed to cause full spectral averaging.

A number of years ago, one of the authors took advantage of this difference in ^1H and ^{13}C “time scales” to resolve a dispute in the literature concerning the site of protonation of amides in strong ($\ll \text{pH } 1$) acid solutions. While it had previously been generally accepted that the carbonyl oxygen was the site of amide protonation in strong acids, *Liler* argued that, in very strong sulfuric acid solutions, there was switch-over to N-protonation (91). This was based on the 60-MHz ^1H spectrum of *N,N*-dimethylformamide in these solutions. This showed two methyl ^1H peaks in both neutral and weakly acidic solutions, due to hindered rotation about the central C–N bond. However, the two methyl signals collapsed to a singlet as acidity increased. *Liler* believed that this indicated predominant N-protonation at high acidity since this would lower the barrier to C–N rotation by minimizing C–N double bond character (91). Since this conclusion was doubted, the ^{13}C spectrum of *N,N*-dimethylformamide was recorded under the same conditions (92). These spectra showed two methyl ^{13}C signals at all acidities, although the signals did broaden to a limited extent at intermediate acidity values. This indicated predominant O-protonation at all acidities, but with a small proportion of N-protonation at intermediate acidities and with fast exchange between the two tautomeric forms. This small fraction of N-protonation allowed slightly faster rotation about the C–N bond that was sufficient to coalesce the slightly separated (*ca.* 6 Hz at 60 MHz) ^1H methyl signals but not the more widely separated ^{13}C signals, which instead were only slightly broadened (92).

When considering exchange processes, one can define three exchange regimes: the slow exchange regime where the frequency of exchange is significantly lower than the peak separation, the intermediate exchange region where the two values are comparable in magnitude, and the fast exchange regime where the exchange rate is much larger than the frequency difference. In the slow exchange region, one will observe separate, relatively sharp, spectra for the two (or more) different forms. In the fast exchange regime, one will observe a single, relatively sharp, spectrum corresponding to the weight average of the spectra for the different exchanging forms. Finally, the appearance of the spectrum of a compound in the intermediate exchange regime will depend strongly on the relationship between the frequency separations between individual pairs of exchanging peaks and the exchange rate. Thus, pairs of peaks with a small frequency separation may appear as a broadened single peak while a pair with a larger chemical shift difference may appear as two broadened peaks. A further complication occurs when the relative populations of the

exchanging forms are significantly different. Consider the exchange between two forms, A and B, in relative proportions of 10:1. The back exchange rate B→A will be 10 times as great as A→B, and peaks for form B will broaden 10 times as fast as the corresponding A peaks with the onset of exchange. Under these circumstances, the minor form may not be observed, particularly if signal/noise is marginal.

Dynamic effects on NMR spectra can create different kinds of problems for a natural product chemist, depending upon which exchange regime is involved. One common problem is very fast interconversion of two or more conformations of a molecule. Provided that the exchange rate is sufficiently fast, the researcher may not be aware that conformational averaging is occurring since the spectrum will not be different in overall appearance from that of a molecule with a fixed conformation. However, both chemical shifts and ^1H - ^1H coupling constants will be the weight averages of the values for the different conformers. Since one often relies heavily on vicinal ^1H - ^1H coupling constants in determining stereochemistry, this could result in misleading conclusions. Furthermore, interproton distances also vary with conformation, so the observed NOE between a pair of protons will also be an average of the NOEs for the different conformations. However, since NOEs vary as r^{-6} (59), the observed NOE will not be a simple weight average but will be strongly biased towards the NOE of the conformer with the shortest interproton distance, even if this is a minor conformer (68b). Thus, NOE data in conformationally mobile systems can be highly misleading if not interpreted with care.

Situations where low barriers are probable to interconversion between conformations of similar energy include molecules with five-membered rings, six-membered rings with fused *cis*-ring junctions, and larger macrocyclic molecules. Unfortunately, there is no easy solution to this problem. If a molecule is in the fast exchange regime at room temperature, the interconversion barrier must be low. Consequently, it may not always be possible to slow the exchange by cooling the solution to the point where well-resolved spectra for individual conformers can be obtained and coupling constants determined, even on a high-field spectrometer. An alternative is to use either molecular mechanics or quantum mechanics calculations to estimate the 3-dimensional structures and relative energies of different significantly populated conformers. A relationship such as the *Altom* equation (60) can then be used to predict the vicinal ^1H - ^1H couplings in the different conformers and determine whether their calculated weight average values are consistent with the observed values.

Systems in the intermediate exchange region present entirely different problems. The biggest risk is that some signals may be so broad that they are not clearly observed. This is most likely to be a problem for ^{13}C spectra because the typically much larger ^{13}C chemical shift differences make extreme broadening more probable. Oddly, this is a situation where the use of a higher field spectrometer can actually be a disadvantage because the larger frequency difference between exchanging signals makes extreme broadening more likely. A second factor is that the lower sensitivity of ^{13}C spectra makes it more likely that a broadened peak cannot be clearly distinguished from noise. One approach that we find effective in cases where this problem is suspected is to reprocess the ^{13}C spectrum with extreme line broadening (e.g. 25 Hz). As illustrated in Fig. 12, this aids in distinguishing broad peaks

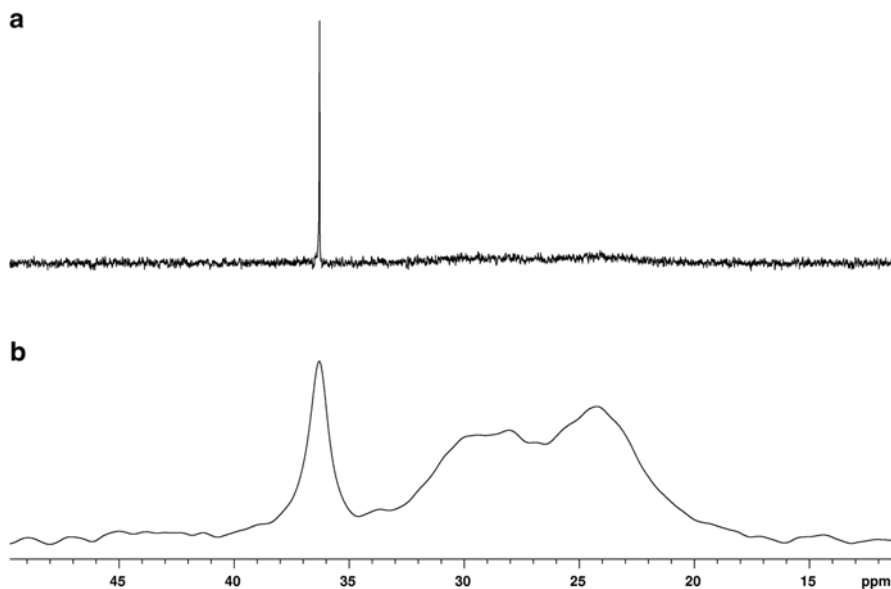
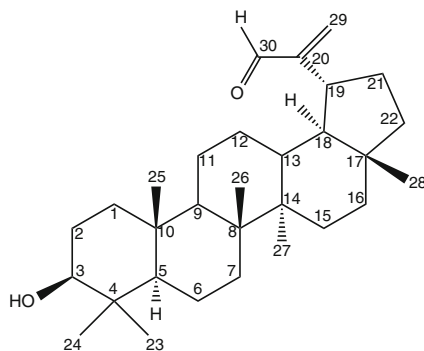


Fig. 12 125-MHz ^{13}C spectrum of *cis*-decalin at 25 °C: (a) with 1-Hz line broadening (b) with 100-Hz line broadening. This illustrates how severe line broadening allows one to detect peaks that are severely broadened by an intermediate exchange rate

from noise. In addition, due to a smaller frequency separation, the proton bonded to the broadened ^{13}C peak may be much sharper. In this case, it is sometimes possible to detect a correlation between a directly bonded $^1\text{H}/^{13}\text{C}$ pair in an HSQC spectrum, which will allow one to determine the ^{13}C chemical shift with adequate precision (93). Similar correlations between indirectly bonded $^1\text{H}/^{13}\text{C}$ pairs may be observed in an HMBC spectrum, although the lower sensitivity of the latter spectrum may make this less likely. Finally, one can repeat the measurements at higher or lower temperatures to attempt, respectively, to move the system to the fast or slow exchange regime. Heating the sample will sharpen the broadened peaks and make them easier to detect. Cooling the sample will potentially allow one to detect and identify the two (or more) exchanging conformations or tautomers. However, this will generally require a significant lowering of temperature to slow the exchange sufficiently to allow sharp peaks to be observed. The choice of which approach to use will also be determined by the liquid range of the solvent used, relative to room temperature. For example, C_6D_6 and, particularly, $\text{DMSO}-d_6$ are suitable for high temperature measurements but freeze not far below room temperature. On the other hand, CDCl_3 and CD_3OD are good for low temperature measurements but of limited value for high temperature measurements due to relatively low boiling points.

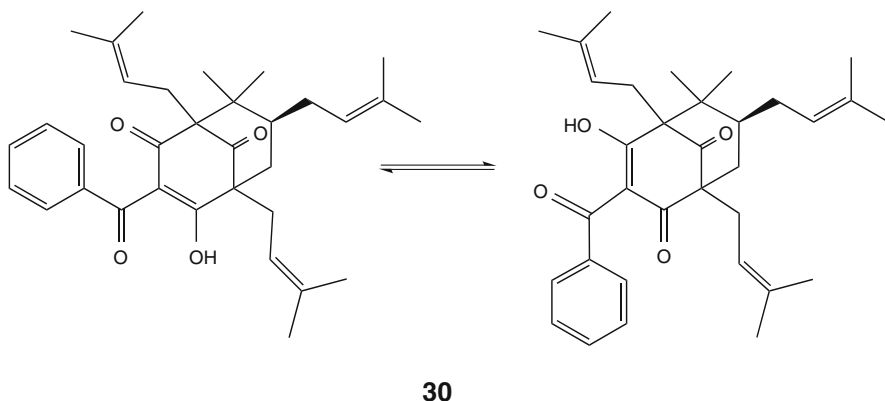
One compound, which nicely illustrates many of these problems (and solutions), is lupane- 3β -ol-30-al (**29**) (93). The original ^{13}C spectrum appeared to show only 25 peaks, which suggested a sesterterpene or possibly a degraded steroid. However, the

^1H spectrum, which showed several methyl singlets, seemed more consistent with a triterpene or possibly a tetranortriterpene. Repeating the ^{13}C spectrum at $-40\text{ }^\circ\text{C}$ revealed four additional peaks, which were between 20 and 40 Hz wide at half-height. Finally, the HSQC and HMBC spectra revealed correlations to another carbon for which the line was still too broad to be clearly observed even in the low temperature ^{13}C spectrum. The NMR data, in combination with molecular modeling calculations, revealed that the observed dynamic effects were due to slow interconversion of two conformations of the side chain aldehyde group (93).



29

In the slow exchange limit, it may not be clear initially whether one is dealing with two interconverting forms or a mixture of two compounds of relatively similar structure. Here, the best approach is to rely on an EXSY spectrum to distinguish between these possibilities (65). This spectrum will show symmetric off-diagonal peaks between the corresponding protons in two exchanging forms. An EXSY spectrum can be obtained with the same pulse sequence that is used for obtaining either a NOESY or ROESY spectrum (65). However, the difference is that EXSY peaks have the same phase as the diagonal peaks while NOESY peaks (for small molecules) and ROESY peaks (always) are of opposite phase to the diagonal peaks. For this reason, NOESY and ROESY spectra should always be obtained in the phase-sensitive mode so that exchange peaks can be distinguished from NOE peaks. Another application of EXSY correlations in the natural product area is the detection of OH peaks hidden beneath other proton peaks. The OH peaks often show EXSY peaks with residual water in the solvent and can be detected by taking a cross-section through the water peak in the NOESY or ROESY spectrum. With the aid of EXSY spectra in combination with other 2D spectra, we find that it is possible to totally assign the structures and spectra of two interconverting forms of even complex natural products. An example is a prenylated benzophenone **30**, where the two tautomeric forms were fully assigned in this way (94). The main problem with this approach occurs when one of the forms is present in only a minor amount since, as noted at the beginning of this section, the minor component peaks may be severely broadened.



10 The Relative Advantages and Disadvantages of Different Pulse Sequences

This topic was extensively discussed in a 2002 review article (10), and the conclusions from that article will be only briefly summarized here. Rather, we will focus mainly on developments since that time. Two key choices, which were previously discussed in the earlier review, were between HMQC and HSQC for one-bond ^1H - ^{13}C correlations and between NOESY and ROESY for investigating NOEs. Our recommendations were to use HSQC in preference to HMQC and ROESY in preference to NOESY. More recent improvements in pulse sequences and hardware support these recommendations, as discussed below.

The original argument for favoring HMQC over HSQC is that the latter requires more pulses and particularly 180° ^{13}C pulses. Therefore, the latter sequences would be prone to poor performance due to incorrect probe tuning, inhomogeneous RF pulses, and incomplete inversion by ^{13}C 180-pulses over the entire spectral window on high-field spectrometers. However, HMQC has the disadvantage that ^1H - ^1H coupling appears along both $F1$ and $F2$ axes while only along $F2$ in HSQC. This yields a sensitivity and ^{13}C resolution advantage for HSQC (95). In addition, HSQC can be run in a phase-sensitive, edited mode (see below) while HMQC is usually run in an absolute-value (magnitude) mode and cannot provide edited spectra. Furthermore, the availability of automatic probe tuning eliminates the first concern about HSQC while modern probe designs now give better pulse homogeneity. Another important advance has been the replacement of “hard” ^{13}C 180-pulses by frequency-swept adiabatic pulses with much greater inversion efficiency (96). Adiabatic pulses also provide more efficient ^{13}C decoupling (96), allowing one to increase the acquisition time and the ^1H resolution without concerns about decoupler heating. This does not require an increased total experiment time since the relaxation delay can be correspondingly decreased to keep the acquisition time constant (10).

A particularly useful version of HSQC is the edited version, which gives peaks of opposite phase for CH_2 carbons relative to CH and CH_3 carbons (97). This provides

the same information as an edited ^{13}C DEPT experiment (10) in comparable time, with the important added advantage of providing the chemical shifts of attached protons (10). However, it has the disadvantage that peaks near the outer edges of the spectral window may be severely attenuated due to a mismatch between the average value of the ^1H - ^{13}C coupling used to calculate delays and the actual coupling for that CH_n pair. A clever approach to this problem was the design of an adiabatic pulse (CRISIS), which took advantage of the approximate linear relationship between ^{13}C chemical shifts and ^1H - ^{13}C coupling constants to minimize this problem (98). A more recent improvement includes a modified CRISIS refocusing pulse during the evolution period with broadband ^1H and ^{13}C inversion pulses during the INEPT and reverse-INEPT stages. The same sequence also provides further sensitivity enhancement by simultaneous acquisition of the two coherence pathways (99). These improvements provide a much more robust version of HSQC, as can be seen in Fig. 13, which shows spectra obtained with the basic gradient HSQC sequence and

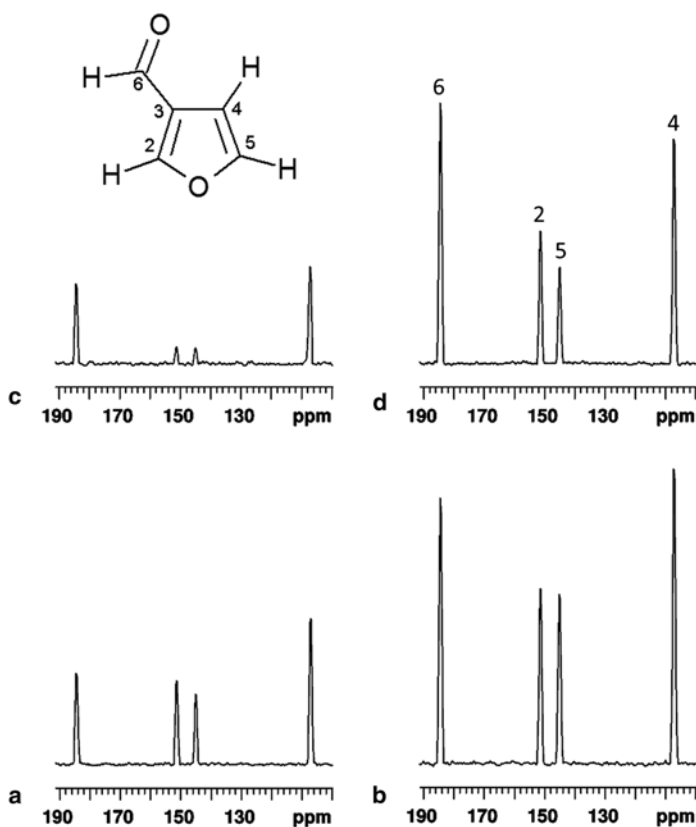


Fig. 13 “Skyline” projection spectra for 3-furaldehyde, using different versions of gradient-selected HSQC pulse sequences: (a) unedited spectrum with the basic gHSQC sequence, (b) unedited spectrum with an improved gHSQC (Agilent gc2hsqcse) sequence, (c) edited spectrum with the basic sequence, (d) edited spectrum with the improved sequence. Carbon numbers are shown at the top of spectrum (d)

one with all of the recent improvements. There are still some sensitivity losses with editing, but these are not nearly as great as with the original sequence.

Recent developments of very fast HMQC sequences (17–19) make them an attractive alternative to HSQC, particularly for rapid screening and dereplication (see Sect. 2). However, these give poorer ^{13}C resolution and cannot provide edited spectra.

Since the wide ^{13}C spectral window is the time-incremented axis, ^{13}C resolution may still be a problem with HSQC (and even more so HMQC), even with the aid of linear prediction. If one is not sample-limited (or if one is fortunate to have access to a ^{13}C -optimized cryogenically cooled probe), it is worth considering acquiring a HETCOR spectrum (35) in cases of severe ^{13}C spectral crowding. As we have shown (100), this can give resolution of close-spaced peaks, which is not possible with HSQC.

There have been numerous proposed modifications of the basic HMBC spectrum, and these have recently been extensively reviewed (101, 102). Unfortunately, in an attempt to improve information content, most of these involve some loss of sensitivity from what is already a low-sensitivity experiment. In contrast, we have shown that one can potentially obtain significant sensitivity enhancements of HMBC spectra by correct choices of acquisition and processing parameters (12). This is illustrated in Fig. 14, which shows a comparison of HMBC spectra of strychnine (1) obtained using our recommended parameters (12) and those from a

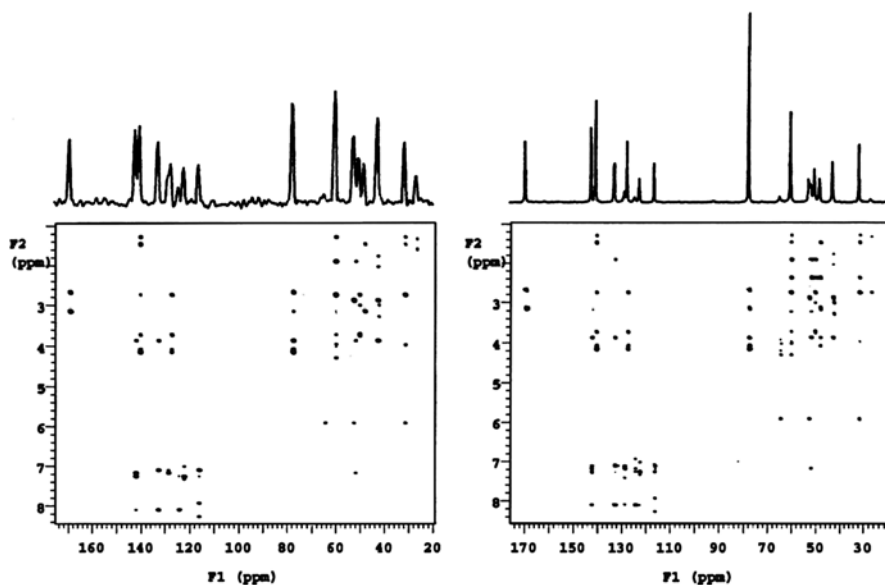


Fig. 14 Absolute-value mode gHMBC spectra of kauradienoic acid (3) with summed projection spectra along the *top*. The *left-hand spectrum* was obtained using suggested acquisition and processing parameters from (8). The *right-hand spectrum* was obtained using recommended parameters from (12). The summed spectrum on the *left* is plotted at 10 times the vertical scale of the summed spectrum on the *right*, and the two summed spectra have respective signal/noise of 22:1 and 150:1. This clearly illustrates the importance of parameter choices in obtaining 2D spectra. This figure was taken from (12) with permission of the publishers

widely used book, which suggests acquisition and processing parameters for numerous 1D and 2D experiments (8). The improvement in S/N , shown in Fig. 14, is actually greater than one would obtain by switching from an ambient temperature probe to a cryogenically cooled probe. In addition, the basic HMBC sequence can be further improved to a limited degree by the incorporation of adiabatic pulses (103).

The types of HMBC modifications that have potential advantages fall into three main categories. The first is a group of sequences, which can separate 2-bond and 3-bond C–H correlations. Of these, the most widely used is the H2BC sequence (104). This relies on the presence of vicinal ^1H – ^1H couplings to generate only ^1H – ^{12}C – ^{13}C correlations. The 2D display produced is similar to that for an HMBC spectrum. Thus, from a side-by-side comparison of the two spectra, one can directly distinguish between 2-bond and 3-bond correlations since only the latter will appear in the HMBC spectrum. However, there are three disadvantages. First, since the correlation information is relayed *via* the vicinal couplings, it does not generate any correlations involving non-protonated carbons, and thus 2-bond and 3-bond correlations to these carbons cannot be distinguished. Second, it requires acquisition of an additional, relatively low-sensitivity, spectrum. Finally, it does not contain any information that could not be deduced from a COSY spectrum or, in case of spectral crowding, from the combination of COSY and HSQC spectra. This suggests an alternative to H2BC. Instead, since one normally would have acquired both COSY and HSQC spectra, covariance processing (45) could be used to generate an HSQC-COSY spectrum from these spectra. Alternatively, one could directly generate an HSQC-COSY spectrum, but this again would require acquiring an additional spectrum. This illustrates what we regard as the biggest advantage of covariance processing, *i.e.* the ability to use two existing high-sensitivity spectra to generate a new spectrum, which would otherwise require significant additional spectrometer time.

One problem with the HMBC sequence is that it uses a fixed delay to generate correlations. This typically is chosen to be optimum for 8-Hz ^1H – ^{13}C couplings. However, this may give very weak correlation peaks in cases where the actual coupling is significantly different from 8 Hz. The second group of modifications are those designed to sample a wider range of long-range ^1H – ^{13}C couplings to minimize this problem. One example is the ACCORD sequence (105), which uses the “accordion” approach to sample couplings over a range as large as 2–25 Hz. This clearly generates a wider range of correlations with good sensitivity (105) but at the cost of introducing a new problem. HMBC spectra, like HMQC, have ^1H – ^1H coupling appearing along both $F1$ and $F2$, generating skewed cross-peaks. The accordion section in ACCORD significantly increases the width of this skew pattern along $F1$, with resultant loss of resolution along that axis. In an attempt to maintain the advantages of ACCORD, while minimizing the skewing problem, *Krishnamurthy* and *Martin* developed the CIGAR sequence (106). This uses a modified accordion section, which can be adjusted to totally eliminate the cross-peak skew, producing an HMBC spectrum with no ^1H – ^1H coupling along $F1$ (similar to HSQC). This can produce significantly improved resolution in regions of spectral crowding (10). However, it also introduces extra delays, which can significantly decrease S/N , particularly for larger (>400 molecular weight) molecules with shorter relaxation

times. Fortunately, F_1 spectral crowding in natural product HMBC spectra is often restricted to a relatively narrow region of the spectrum. While one could run a band-selective HMBC spectrum for this region, an alternative would be to obtain a band-selective CIGAR spectrum. Other approaches to sampling a wide range of couplings include converting HMBC to a 3D experiment with J_{CH} forming the third axis or by combining the data from three or four experiments with delays corresponding to different values of J_{CH} . The advantages and problems of these approaches are discussed elsewhere (101, 102).

The third, and in our view most promising, HMBC modification is the IMPACT-HMBC sequence, which was recently developed by *Furrer* (107). This is similar in design to the ASAP-HMQC sequence of *Kupce* and *Freeman* (18) in that it uses cross-polarization of protons to allow a far shorter (*ca.* 0.2 s) relaxation delay. This either permits one to acquire spectra more quickly or to collect more scans per time increment in a given time, increasing the signal/noise.

We have previously argued for the use of ROESY in place of NOESY for NOE investigations (10). The main advantage of ROESY is that cross-peak intensities are nearly independent of molecular weight while NOESY cross-peaks change sign as molecular weight increases. Depending on solvent viscosity, the crossover point typically occurs somewhere in the 500–1,500 molecular-weight region. Consequently, particularly for larger natural products, NOESY cross peaks may be very small. However, one problem with ROESY has been that the spin-lock generates heat if it is too long. Recent improvements replace the original spin-lock, which used hard pulses with a lower power adiabatic pulse spin-lock that allows the use of longer mixing times (108).

Finally, one disadvantage of the original TOCSY sequence (34) was phase distortions, which altered the appearance of cross-peaks. This has been minimized by the more recent Z-TOCSY experiment, which incorporates a zero-quantum filter, yielding cleaner spectra and well-phased peaks (67). These can also be incorporated in the selective 1D TOCSY sequence. An example of this was given in Sect. 7.

11 Liquid-Chromatography–NMR

Since high-pressure liquid chromatography (HPLC) is so widely used for isolating pure natural products from chromatographic fractions or other complex mixtures, the combination of LC with NMR would seem to be a logical approach to use in natural product research. This approach has been investigated widely over a number of years (109). However, the relatively low sensitivity of NMR has been a persistent problem when trying to use continuous flow LC in combination with NMR. The conditions for optimum resolution of LC peaks usually leave too little sample in the flow cell of the NMR probe during acquisition to allow one to obtain anything more than a routine ^1H spectrum. While this problem can be partially overcome if one is fortunate enough to have access to a cryogenically cooled flow probe (see Sect. 12), the consensus seems to be that it is better to use the two techniques separately.

One intermediate approach would be to use stop flow LC for sample collection. However, most workers in the field seem to agree that the best approach is to use solid phase extraction (SPE) cartridges to collect the LC fractions (109). The samples can then be dissolved off the cartridges using deuterated solvents and either injected into a flow NMR probe or placed in NMR tubes to be used in conjunction with a sample changer and a regular NMR probe. There are two major advantages to this approach. First, the LC separation can be carried out with protonated solvents, minimizing costs. Second, one can use repeat injections to increase the amounts of samples collected, if necessary.

A further modification of this approach is to combine LC, mass spectrometry (MS), and NMR (110). Liquid chromatography is used to separate a complex mixture into individual components, using protonated solvents. As each component is detected, usually by a UV-visible detector, the fraction concerned is split, with a small amount sent for MS analysis, with the remainder sent to an SPE cartridge for collection. If the MS does not allow identification of the fraction as a known compound, the solvent can be removed from the SPE cartridge and the sample reconstituted in the appropriate deuterated solvent for NMR investigation. Although neither of the authors has extensive experience of combining LC with mass and NMR spectroscopic analysis, we believe that this is a promising approach since it combines dereplication (*i.e.* distinguishing known from unknown compounds) with full structure determination, when the latter is required.

12 Probe Choices

In most cases, a natural product chemist may have limited probe choices available, as determined by the available probes in the NMR facility. However, it is still useful to have a general knowledge of the relative advantages and disadvantages of different probe types. These are discussed below.

12.1 *Essential Probe Features for Natural Product Research*

Any probe should have H/X capabilities, *i.e.* observation of both proton and heteroatoms (most commonly ^{13}C but preferably also at least ^{15}N). This can involve either an H-channel and a tunable X-channel or a three-channel (H/C/N) probe with separate channels tuned to the three nuclei. It should also have a z-axis gradient coil for gradient shimming and performing gradient-selected 2D NMR sequences. Finally, it is highly desirable to have auto-tuning capabilities for both H and X channels. This is particularly important if the spectrometer is being operated with an auto-sampler but also for multi-pulse experiments (*e.g.* 2D NMR) where having one or both channels out of tune can significantly degrade performance and introduce artifact peaks.

12.2 *Ambient-Temperature Probes*

The main advantages of ambient-temperature probes are their low capital and operating costs, but they are significantly less sensitive than the alternative cryogenically cooled probes (see below). They most commonly have inserts for either 5-mm or 3-mm sample tubes. A key feature, particularly in the past, has been the geometry of the two coils in an H/X probe since the inner coil was relatively more sensitive than the outer coil (often by a significant extent). If the H-coil is inside, this is commonly called an indirect-detection probe while, if the X-coil is inside, it is called a direct-detection probe. This odd terminology is a historical one, dating to the time when 2D H/X experiments almost always involved X-detection. The difference in coil sensitivities has been a problem for natural product chemists since one usually wished to obtain both a 1D ^{13}C spectrum and a series of ^1H -detected 2D spectra on the same sample, preferably without having to change probes. Fortunately, some of the latest generation of probes from both Bruker and Agilent have minimized this problem by providing good sensitivity on both coils. For example, a probe of this type, to which both authors of this chapter have access, is the Agilent “OneNMR” probe that gives ^1H and ^{13}C *S/N* specifications, which are, respectively, almost identical with the corresponding specifications for ^1H on an indirect-detection probe and ^{13}C on a direct-detection probe from the same manufacturer.

12.3 *Cryogenically Cooled Probes*

With these probes, the coils and the preamplifiers are cryogenically cooled, usually with liquid He, but with the sample at ambient temperature. The actual coil temperature is usually about 20 K. This very significantly reduces random thermal noise, leading to a dramatic increase in *S/N* compared to ambient-temperature probes. The enhancement factor is usually quoted as about 4:1, but the actual *S/N* enhancement appears to be strongly solvent-dependent with larger than 4:1 enhancements for some common organic solvents (*e.g.* CDCl_3 and C_6D_6 , in particular) and less than 4:1 for aqueous solutions, particularly “salty” solutions. While the sensitivity of these probes is obviously a major advantage, the disadvantage is that they are not only far more costly than regular probes but also require expensive routine maintenance every 1–2 years. They also appear to be more prone to other damage and are costly and time-consuming to repair. In times of shrinking budgets, one must carefully balance the sensitivity advantages against the significant maintenance costs in deciding whether to acquire this type of probe. If one does have sufficient funds to include a cryogenically cooled probe in a spectrometer purchase, one should also consider the alternative of extra ambient-temperature probes plus a significantly extended warranty on the entire spectrometer package. The most common cryogenically cooled probes are indirect-detection H/C/N probes. While these were designed mainly for protein NMR studies, they

are also quite suitable for natural product investigations. Alternatively, both major manufacturers offer high sensitivity ^{13}C -optimized probes, which still give ^1H S/N specifications well in excess of ambient-temperature indirect-detection probes. Arguably, these would be a better choice for natural product research since they can quickly provide the good quality ^{13}C spectra, which are required by journal editors for publication, while still providing excellent sensitivity for ^1H -detected 2D experiments.

Cryogenically cooled probes are most commonly designed for 5-mm or 3-mm tubes although Bruker also offers an H/C/N 1.7-mm probe. One point to remember is that, with a 5-mm probe, there are often S/N advantages in using 3-mm tubes (111). The reason is that, with the cryogenic cooling of coils and preamplifier, the main source of thermal noise is the sample. Thus, providing that solubility is not an issue, there is an advantage in reducing the sample volume with a 3-mm tube.

Bruker has recently offered indirect-detection liquid N_2 -cooled H/X probes for 400–600 MHz instruments. While the S/N is only about half of that of the LHe-cooled probes, capital costs are lower. However, they require continuous cooling with liquid N_2 , so operating costs will be higher than for ambient probes.

Finally, the ultimate cryogenically cooled probe for natural product research is one built specially for the National High Field Magnet Laboratory in Florida (112). This not only has cooling of the coils and preamplifiers but also has coils fabricated from superconducting materials. It has been used to elucidate structures of marine natural products at the nanomole level (113). While there are rumors of a possible commercial version of this probe in the future, it would undoubtedly be considerably more expensive than current cryogenically cooled probes.

12.4 Microprobes

Microprobes represent a different approach to probe design, which is particularly suitable for sample-limited cases, namely, to have a very small sample volume. A commercial version of a probe of this type is the Protasis CapNMR probe, which is compatible with spectrometers from all major manufacturers. The probe requires about $15\text{ m}\mu^3$ of solution with an active volume (in the form of a flow cell within the probe) of $5\text{ m}\mu^3$. This is available as an H/C probe, either with a single-flow cell plus gradient coils or with dual-flow cells without gradients. The latter arrangement allows for parallel acquisition of spectra from two samples. Samples can be loaded with a robot auto-sampler, allowing for high throughput operation. The main limitation would appear to be sample solubility. Nevertheless, CapNMR does provide an intermediate cost alternative to cryogenically cooled probes and has proved useful in natural product research, particularly when used in combination with HPLC-SPE separation techniques (114).

13 A Fully Automated Setup of 2D NMR Experiments for Organic Structure Determination

There have been a number of advances in spectrometer operation, which minimize the extent of operator interaction with the spectrometer. These include robots for sample changing, automated locking and probe tuning, gradient shimming and improved software for experiment setup. However, with the increasing speed and power of computers, we believe there is still room for further significant improvement, particularly in automated setup of acquisition and processing parameters for 2D experiments. This would allow replacing the default parameters currently included with the spectrometer software for different pulse sequences with parameters optimized for the actual sample, without requiring expert knowledge on the part of the operator. A possible future program to fully achieve these goals is outlined below. However, in the interim, some of the ideas could be quickly implemented, *e.g.* using a quick ^1H T_1 measurement to choose optimum recycle times for different experiments.

The program could provide a menu of standard 2D experiments (probably at least COSY-45 or COSY-90, NOESY/ROESY, TOCSY, HSQC, and HMBC), along with a Help file indicating the information content and the relative sensitivity of different experiments. It could also provide the option of inputting an estimated ^{13}C spectrum (calculated with existing 3rd party software), provided that the probable structure was known with reasonable certainty. The first step would be to specify the 2D experiments to be run and also whether the operator wished to obtain a DEPT-135 or DEPT-Q ^{13}C spectrum. The spectrometer would then be instructed to acquire a proton spectrum with a default number of scans (probably 16) and with a wide spectral window to ensure that no peaks are missed. The spectrometer could then reacquire the spectrum with the spectral width narrowed to include only regions where peaks appeared and the number of scans adjusted to the minimum number needed to give good signal/noise. The next step would be to have the spectrometer measure proton T_1 values by finding nulls in a quick inversion-recovery experiment. It could be programmed to ignore the most intense peaks (solvent peaks and methyl signals) so that T_1 values were determined only from the weaker CH and CH_2 multiplets (CH_3 signals usually have longer relaxation times but are also much more intense. Thus, one can afford to have a shorter than optimum recycle time for these protons).

Based on the measured signal/noise and number of scans for the ^1H spectrum and the known relative ^1H and ^{13}C sensitivity of the probe in use, the program could then calculate the time needed to run a 1D ^{13}C spectrum or else either a DEPT-135 or a DEPT-Q spectrum (the former giving only peaks for protonated carbons while the latter includes all types of carbons but with significantly reduced sensitivity for quaternary carbons). Next, it could optimize the acquisition parameters for the chosen 2D experiments. The minimum number of data points required for the acquisition axis depends on the extent of ^1H spectral crowding. One way to estimate this would be to use a binning technique for the proton spectrum similar to that used in metabonomics investigations by NMR, *i.e.* the spectrum could be divided into a

series of “bins” of equal frequency width (maybe 20 Hz) and integrated. The density of peaks could be estimated from the fraction of bins with significant integrated area and/or the number of consecutive bins with significant area. This would allow the computer to choose the minimum number of F_2 data points and acquisition times for all 2D experiments. The recycle time (acquisition time plus relaxation delay) would be set at 1.3 times the average T_1 determined above for most experiments except for NOESY or ROESY where 2.5 times T_1 would be more appropriate. The minimum number of required data points determined by binning could then also be used as the number of F_1 time increments for homonuclear 2D experiments (the number of time increments should include both acquired and linearly predicted increments to minimize total time).

If the operator wished to include heteronuclear experiments (*e.g.* HSQC and HMBC), the program could again use the number of scans and the signal/noise for the initial proton spectrum plus the known relative ^1H and ^{13}C sensitivities for the installed probe to calculate the number of scans needed to acquire these experiments. If a calculated ^{13}C spectrum were available, then the HSQC and HMBC ^{13}C spectral windows could initially be chosen based on this spectrum. Alternatively, it could initially choose a default value of 225 ppm for HMBC and a default value for HSQC of 170 ppm or it could scan the initial ^1H spectrum and choose the latter spectral window based on the presence or absence of peaks in regions characteristic of aromatic/olefinic and aldehyde protons. If a calculated ^{13}C spectrum were available, then the program could use average peak separations to determine the minimum number of measured and linearly predicted time increment spectra needed to get adequately resolved HSQC and HMBC spectra. Otherwise, it could use the extent of proton spectral crowding to estimate the number of required time increments.

In addition, the program would then list the time for each experiment (including the alternative 1D ^{13}C options) and the total time. The operator could then choose to instruct the spectrometer to proceed with the full set of experiments or, if the calculated time exceeded the available time on the spectrometer, eliminate one or more experiments from the queue. Assuming that either a DEPT-135 spectrum and a full ^{13}C spectrum or else only a DEPT-Q spectrum were run first, the ideal arrangement would be for the spectrometer to have on-line access to a large 1D ^{13}C data library of known compounds. If a close match to a known compound were found, then 2D acquisition could be automatically aborted unless the operator had indicated that it should proceed even if a match were found. Finally, if HSQC and HMBC spectra were to be acquired, the program could first re-optimize the ^{13}C windows for these experiments, based on the ^{13}C spectra if these had been obtained. The same approach could be used for multiple samples in an overnight or weekend run. There could be a “multiple sample” option. In this case, the spectrometer would be programmed to sequentially run the ^1H setup experiments on each sample and then list the times for each sample. The operator could then choose to delete individual experiments, or entire samples from the queue, if the total time were too long.

After completion of data acquisition, the spectrometer could be programmed to process the spectra, based on the pre-determined best weighting function for each axis for each experiment, with the value of the weighting function based on the number of points/number of increments and the chemical shift window. While the

whole procedure may seem cumbersome, it actually closely mirrors the thought processes that a highly experienced operator would use in setting up a series of 2D experiments for organic structure elucidation, while avoiding the risk of operator error. It is also easily within the capabilities of current high-speed computers and would require a minimum of calculation time. Finally, the availability of an artificial intelligence software program of this type would, in combination with automated probe tuning and gradient shimming, allow even an inexperienced operator to acquire a high quality set of ^1H , ^{13}C , and 2D NMR spectra necessary for organic structure elucidation in the minimum possible time.

While a fully automated program of this kind would be ideal, there are also partial steps that could easily be incorporated into current spectrometer software, which would improve the ability of an inexperienced operator to obtain good quality spectra in minimum acquisition times. Software already exists that allows one to estimate the times needed to obtain different 2D spectra, based on the S/N for a 1D proton spectrum. An automated program for ^1H T_1 measurements could be added, with the results used to calculate optimum recycle times for 2D experiments, thus avoiding the common problem of wasted time due to unnecessarily long relaxation delays (10).

14 Parameter Choices for Acquisition and Processing of 1D and 2D NMR Spectra

We include in this section some of the basic background, which explains the reasons for various parameter choices. Sections 14.1 and 14.3 are recommended reading for anyone who likes to understand why some parameter choices are better than others. For those who just want to have simple “menus” for acquiring spectra, Sects. 14.2 and 14.4 partially satisfy this need. However, even then, it is still essential to make some parameter choices, depending on sample amount, molecular weight and the extent of spectral crowding, in order to get the best quality spectra in the shortest possible time. Therefore, we will give ranges of values for key parameters for each type of spectrum, briefly indicating how to choose the most appropriate values. Simply relying on one standard data set for each type of experiment, regardless of the nature of the compound being, will often yield inadequate results.

14.1 Basics of NMR Data Acquisition

14.1.1 Sampling Rate

The *Nyquist* theorem tells us that to define a spectral window that is N Hz wide, we must sample the data at a rate of $2N$ data points per second. The actual number of data points collected will depend on the acquisition time, which is typically of the order of 1–5 s in 1D NMR but shorter in 2D NMR.

14.1.2 Analog to Digital Conversion

The signal detected in the NMR receiver is in continuous (analog) form and must be converted to a digital format for data storage and processing. This is done with an Analog-to-Digital Converter (ADC). The ADC has two key characteristics. The first is the maximum speed of the ADC in Hz, which in turn determines the maximum spectral window that can be determined ($1/2$ of the maximum sampling speed). The second is the binary bit length, which determines the dynamic range of the ADC, *i.e.* the ability to detect weak signals in the presence of strong signals. Until recently, maximum sampling rates of ADCs were typically in the range of 100–500 KHz while a typical ADC had 16 bits. With one bit used to determine the sign, the remaining 15 bits provided a theoretical dynamic range of 32,768:1. However, since it takes 2–3 bits to define a weak peak with reasonable precision, the effective dynamic range was less. In addition, it was critical to adjust the receiver gain so that the detected signal almost filled the ADC in order to achieve this dynamic range. However, the latest model NMR spectrometers have much faster ADCs (up to 80 MHz). This has allowed a new approach to data acquisition, called digital oversampling, which dramatically increases the effective dynamic range of the ADC.

14.1.3 Digital Oversampling

With digital oversampling, instead of sampling at a rate of two times the desired spectral width in Hz (as stipulated by the *Nyquist* theorem), one instead samples at the maximum rate of the ADC. For example, consider a situation where the desired spectral window is 5,000 Hz, but instead of sampling at 10 KHz, one sampled at 80 MHz, *i.e.* 8,000 times faster than the nominal rate. Then, each successive block of 8,000 points is summed to produce a single point. The final result would be a collected FID with the appropriate number of data points for the desired spectral window. With modern high-speed computers, this can be done “on the fly” (in 32-bit arithmetic), *i.e.* during the actual data acquisition. In this case, information theory tells us that the dynamic range is increased by the square root of the extent of oversampling, *i.e.* by about 90:1 for oversampling by 8,000. The 80-MHz ADC on the authors’ latest spectrometers actually has a 14-bit ADC, but oversampling effectively converts it to about a 20-bit ADC, *i.e.* a dynamic range of ~500,000:1. However, there is one caveat to this. Since the data are processed in the ADC prior to the averaging process, it will still be necessary in cases of extremely strong solvent signals (*e.g.* H₂O) to use some form of solvent suppression to avoid overload. On the other hand, in the absence of one or more very strong peaks, it is no longer critical to set the gain as carefully at the start of the experiment.

There is one additional advantage to oversampling. The ADC does not distinguish between a signal, which only partially fills a bit, from one that almost completely fills it. This introduces randomness, called digitization noise. The act of summing a large number of points almost totally eliminates this source of noise. This may have a minimal effect when using an ambient-temperature probe, where thermal noise will usually be the predominant noise source. However, it can make a

significant difference for cryogenically cooled probes where cooling the coils and preamplifier to *ca.* 20 K minimizes thermal noise.

14.1.4 Quadrature Detection

The receiver in an NMR spectrometer is actually a phase-sensitive detector, *i.e.* it measures frequencies relative to the transmitter frequency rather than absolute frequencies. A single phase-sensitive detector cannot distinguish between frequencies that are positive or negative with respect to the carrier frequency. In the early days of FT NMR, this problem was avoided by having the transmitter frequency at one end of the spectral window so that all peaks would have the same sign. However, this introduced two other problems. First, this required a more intense transmitter pulse in order to uniformly excite the entire spectral window. Second, noise would be detected at both positive and negative frequencies, and the noise from the other side of the pulse would “fold in” to the spectral window, reducing signal/noise by $\sqrt{2}$. Both of these problems were solved by the technique of quadrature detection. This involves detecting two signals at right angles to each other. When performed, this permits distinction of positive and negative frequencies, allowing one to put the carrier frequency at the mid-point of the spectral window. In the past, this was normally done by splitting the signal and routing it to two phase-sensitive detectors with a phase shift of 90° between them (accomplished by a very slight delay in sending the signal to the second detector). The signals are sent to different ADCs for digitization and then to two separate memory blocks in the computer for storage. The signals are then *Fourier* transformed and added to provide the final spectrum. Older Varian spectrometers used this approach. Newer Varian/Agilent spectrometers differ in that the signal is first digitized with a high speed ADC and then split into two signals phase-shifted by 90° for storage. This allows the use of a single ADC in place of two of these. A different approach is used on older Bruker spectrometers, which also allowed the use of a single ADC. In these cases, data were actually sampled at a frequency of $4N$ Hz, with a 90° phase shift for each successive data point. However, newer model Bruker spectrometers use an approach, which we believe is similar to that used on Agilent spectrometers.

14.1.5 Fold-in Peaks

The collection of digitized data introduces another problem. The RF pulse also excites peaks outside the chosen spectral window. Since one is sampling at finite intervals, it is impossible to distinguish between peaks, which are just outside the spectral window and those just inside it. Thus, the former peaks will also appear within the spectral window. However, their exact positions depend on the form of quadrature detection used. With the older Varian method, peaks which are outside the right hand (low frequency) side of the window by x Hz will appear x Hz inside the left hand end of the spectral window and *vice versa*. On the other hand, with the older Bruker method, peaks outside of either end of the spectral window will appear at an equal

distance inside the same end of the spectral window. It is often possible to distinguish fold-in peaks because they have different phases than other peaks in the spectrum. Fortunately, newer spectrometers from both manufacturers use filters, which are extremely effective at suppressing fold-in peaks so this is no longer a concern.

14.1.6 Analog Versus Digital Filters

In addition to peaks folding in, noise will also fold in from outside the spectral window, degrading signal/noise. To minimize this problem, spectrometers are equipped with audio-frequency filters. Older model spectrometers applied filtration to the analog signal and are thus called analog filters. While they were set to cut off frequencies just outside either end of the spectral window, this cut-off was not very sharp, with the result that intensities of peaks near either end of the spectral window were somewhat attenuated while some fold-in of peaks could also still be observed. Newer model spectrometers all employ digital filters, which have much sharper cut-offs and thus avoid the problems of analog filters. However, because they are so efficient, one must use caution to ensure that the spectral window is wide enough to include all possible peaks. Otherwise, the user will not be aware that these peaks are actually present in the true spectrum.

14.2 Recommended Acquisition and Processing Parameters for 1D Spectra

14.2.1 Spectral Widths

Since one can usually acquire ^1H spectra very quickly, the authors recommend acquiring an initial “scan” spectrum with a very wide spectral window (*ca.* -1.0 to $+15$ ppm) to ensure there are no unexpected peaks with unusual chemical shifts. Then, a second spectrum can be obtained using a spectral window, which is narrowed to include only observed peaks in order to get better resolution. However, if using an older spectrometer with analog filters, one should leave regions (of *ca.* 1 ppm) with no peaks on both sides of the spectral window, particularly if one wants quantitative peak intensities. Since the lower signal/noise of ^{13}C spectra and longer acquisitions will usually make it undesirable to obtain two spectra, it is recommended using a spectral window wide enough (*ca.* -5 to 225 ppm) to include all possible peaks when acquiring ^{13}C spectra.

14.2.2 Number of Data Points and Acquisition Times

The number of data points, NP, will be given by $\text{NP} = 2(\text{SW})(\text{AT})$ where SW is the spectral width and AT is the acquisition time. One typically chooses the number of points to be some power of 2, *e.g.* 32,768 or 65,536 (often abbreviated as 32 K or

64 K), although this is not essential. For a ^1H spectral width of 5,000 Hz, these two values of NP would, respectively, correspond to acquisition times of *ca.* 3 and 6 s. while for a 30,000-Hz ^{13}C spectral width, *AT* would respectively be *ca.* 0.5 and 1 s. *AT* values of 3–6 s will give reasonable data point resolution for ^1H spectra (see Sect. 14.2.4) so either 32 K or 64 K would be an acceptable choice. However, the use of 64 K points is recommended for ^{13}C . Alternatively, if one is setting *AT*, 4–5 s for ^1H and 1 s for ^{13}C are suggested as acceptable values.

14.2.3 Number of Scans (Transients)

With earlier model spectrometers, it was recommended that the number of scans (NS) should be some multiple of four. This was to allow for a four-step phase cycle, which cancelled “quadrature image” peaks. These arose from imperfections in spectrometer hardware and appeared as weak mirror images of very strong peaks in the spectrum. However, now that quadrature detection is carried out on digitized data, quadrature images are non-existent on later generation spectrometers. Thus, ^1H spectra can be obtained with as little as one scan. The exact number of scans will depend on sample concentration and probe sensitivity. ^{13}C spectra will generally require many more scans. If this option is available, use of the “block size” command is strongly recommended. This is set at a value such as 1 or 40, and the data are stored at the end of each block. This allows one to monitor the *S/N* during acquisition and terminate it when *S/N* is satisfactory.

14.2.4 Zero Filling and Data Point Resolution

Fourier transformation of a FID yields both real and imaginary spectra, with half of the data points used to define each spectrum. Thus, the data point resolution (in Hz/point) is given by $2\text{SW}/\text{NP}$ (or $1/\text{AT}$). However, the data point resolution can be improved by a factor of two, simply by adding an equal number of zeros to the end of the digitized FID. This method, called zero filling, effectively allows the use of all of the experimental data points to define the real spectrum. In this case, the data point resolution becomes SW/NP (or $1/2\text{AT}$) Hz/point. This will provide improved spectral resolution, particularly if the natural line widths of spectral peaks are less than the data point resolution. Zero filling by more than a factor of two will not further narrow spectral peaks. However, it will provide better definition of peak frequencies and a cosmetic improvement in the appearance of complex multiplets. For that reason, we strongly recommend using extra zero filling up to at least 4NP or even higher. On Varian/Agilent spectrometers, this is set by the parameter “fn” while on Bruker spectrometers, the corresponding parameter is “si.”

14.2.5 Pulse Widths and Delay Times

For most ^1H spectra obtained using a 4–5 s acquisition time, one can obtain at least semi-quantitative peak areas using 90° pulses and no relaxation delays, although the areas of methyl groups, which typically will have the longest relaxation times, may be partially suppressed. However, if quantitative peak areas are important, then one has the choice of either using a shorter pulse or including a relaxation delay between scans. *Richard Ernst* investigated this problem in the early days of FT NMR and demonstrated that it was better to use a reduced pulse angle rather than a relaxation delay (3). This can be understood by a simple trigonometric argument. If, for example, one chooses a pulse width corresponding to a 45° rotation of the magnetization vector (a 45° pulse), the component of magnetization along the y-axis is $\sim 71\%$. However, the residual component along the z-axis is also $\sim 71\%$ ($\sin\theta = \cos\theta = 0.71$). Thus, it takes significantly less time for magnetization to return to equilibrium along the z-axis. The optimum *Ernst* angle is given by $\cos\theta = \exp(-AT/T_1)$ where T_1 is the relaxation time, which decreases with molecular size. In practice, we find that a 45° pulse plus a 4–5 s acquisition time will yield quantitative results for most typical natural products, other than those of very low molecular weight (<250 Da). In the latter case, a 30° pulse is suggested, possibly along with a relaxation delay.

The choices for ^{13}C spectra are more difficult, because the acquisition times are shorter, and there are typically much wider ranges of relaxation times, with quaternary carbons having the longest values of T_1 . However, due to differences in NOEs for different carbons, ^{13}C spectra are rarely quantitative. Thus, the authors believe a pulse width of 45° , combined with a 1-s relaxation delay, will generally give satisfactory results.

14.2.6 Apodization (Weighting) Functions

With ^1H spectra, it should not usually be necessary to use any form of apodization function, provided that the FID has decayed below the noise level at the end of the acquisition time. However, for spectra with poor *S/N*, a small amount (*ca.* 0.1–0.3 Hz) of exponential line broadening can be used. This will improve signal/noise at the cost of a small loss of resolution. On the other hand, if one wants to improve the resolution of a spectrum, a resolution enhancement function can be used, which combines a positive *Gaussian* function with a negative line broadening function. The right combination of these two parameters can be set by the spectrometer software (for Varian/Agilent spectrometers the command is “*resolv*”). These parameters should be chosen based on twofold zero filling, *e.g.* 32 K to 64 K points. However, after the parameters are chosen, one can further increase the amount of zero filling. This will aid in accurately determining splittings in multiplet patterns. Note, however, that resolution enhancement will degrade signal/noise and that relative peak areas may no longer be quantitative. In particular, if a spectrum has both well-resolved multiplets and broad peaks, the latter will be suppressed.

In the case of ^{13}C spectra, some extent of line broadening is usually needed to improve signal/noise. If there is an interactive weighting program available on a given spectrometer (“wti” on Varian/Agilent spectrometers), the ideal approach is to choose a weighting function having the same decay time as the FID (a “matched filter”). Otherwise, the authors suggest 1–3 Hz line broadening, depending on the signal/noise.

14.2.7 ^{13}C Spectral Editing

When assigning ^{13}C spectra, it is helpful to assign the type of carbon, *i.e.* quaternary, methine, methylene or methyl, to each signal. There are two basic ways of achieving this, with neither entirely satisfactory. The first is to use the APT sequence, which produces peaks for quaternary and methylene carbons, and are of opposite phase to those for methine and methyl carbons (115). This sequence has two disadvantages: it is less sensitive than a regular ^{13}C spectrum and is sensitive to variations in the one-bond ^{13}C – ^1H coupling constants, potentially giving misleading results (10). The second is the DEPT spectrum (116). This involves polarization transfer from directly bonded protons to carbons and can be either used to generate a spectrum (DEPT-135) with methylene carbons of opposite phase to methine and methyl carbons) or, by combining DEPT-45, DEPT-90, and DEPT-135 spectra, to produce separate spectra for the three types of carbons (10). The main advantages of DEPT are that it has better signal/noise than a regular ^{13}C spectrum and is significantly less sensitive than APT to variations in $^1J_{\text{CH}}$. The main disadvantage is that it only gives peaks for protonated carbons and thus it will still be necessary to also record a standard ^{13}C spectrum to observe all carbons. An alternative version of DEPT, called DEPT-Q, has been developed that also shows quaternary carbons (117). However, the quaternary carbon signals are generally weaker than those observed in a standard ^{13}C spectrum obtained in the same time. Nevertheless, it may still be faster to obtain a DEPT-Q spectrum than separate DEPT and standard ^{13}C spectra. Finally, an alternative approach, which the authors favor, is to instead acquire an edited HSQC spectrum. This provides the same information as DEPT in comparable or less time, with the additional major advantage of providing assignments for the directly bonded hydrogen(s) associated with each carbon (10). However, it does not provide as accurate ^{13}C chemical shifts as DEPT and, again, provides no information about quaternary carbons.

The parameter choices for APT spectra will generally be the same as would be used for standard ^{13}C spectra. The suggested value of $^1J_{\text{CH}}$ used to calculate editing delays is 145 Hz. However, an unusual feature of APT is that one should not replace the initial ^{13}C 90° pulse by a 45° pulse. Since there is a later ^{13}C 180° pulse, this will convert the residual *z*-magnetization remaining after the original pulse to $-z$ -magnetization, requiring an even longer relaxation delay. Instead, the initial pulse should be a 135° pulse since, in this case, the initial residual *z*-magnetization will be along the $-z$ -axis but converted to $+z$ -magnetization, decreasing the needed relaxation delay. With this modification, a 1-s relaxation delay should be sufficient.

With DEPT, one is transferring magnetization from ^1H to ^{13}C , so it is the ^1H relaxation time that matters. Since ^1H decoupling is applied during acquisition, the relaxation delay must allow recovery of this magnetization. The recommended delay is $1.3T_1$, so a relaxation delay of 1–1.5 s should be sufficient for typical natural products. Again, a value of $^1J_{\text{CH}}$ of 145 Hz is suggested. Finally, since residual ^{13}C magnetization is cancelled by phase cycling, it is essential to include a series of steady-state (dummy) scans before acquiring data in order to establish steady-state ^{13}C magnetization. Otherwise a residual solvent peak will be observed which may obscure desired peaks. Generally, 16 dummy scans should be sufficient.

14.3 Basics of 2D NMR

14.3.1 General Features of 2D NMR Sequences

2D NMR sequences are generally composed of three components: an initial relaxation delay, a variable evolution period designated as t_1 and usually initiated by a 90° pulse, and finally an acquisition period, usually labeled t_2 . A series of spectra are generated by incrementing the evolution time, usually in a regular fashion. The F_1 spectral window is determined by the number of increments and the time between increments. The latter is automatically calculated by the pulse sequence program based on the specified number of increments and F_1 spectral width. The key differences between different sequences almost all occur during t_1 , in the form of additional fixed delays and/or additional pulses. The FIDs acquired during t_2 for different values of t_1 are *Fourier* transformed to yield a series of F_2 spectra. Then, the phase and intensity of each point in t_2 for the different time-incremented spectra is used to generate a series of t_1 interferograms, which resemble FIDs. These are then *Fourier* transformed to produce the second frequency axis, labeled F_1 . Conventionally, spectra are plotted with F_2 as the horizontal axis and F_1 as the vertical axis.

14.3.2 Homonuclear and Heteronuclear 2D NMR Spectra

Homonuclear spectra have spectral information for the same nucleus (usually ^1H) along both axes. The spectrum has a 1D spectrum along a diagonal from bottom-left to top-right and symmetric off-diagonal peaks between interacting nuclei. COSY or TOCSY spectra are generated when the interaction between different nuclei is due to scalar (“ J ”) coupling while NOESY or ROESY spectra are generated when the interaction is due to dipolar relaxation. Heteronuclear spectra provide information about scalar coupling between heteronuclei (mostly commonly $^1\text{H}/^{13}\text{C}$ but also $^1\text{H}/^{15}\text{N}$ in the natural product area). In this case, the spectrum for the acquired nucleus is plotted on the horizontal axis, while information about the second nucleus (generated by incrementation of t_1) is along the vertical axis. The cross-peaks indicate either one-bond or n-bond ($n=2$ or 3) $^1\text{H}/^{13}\text{C}$ coupling, depending on the sequence used. Early $^1\text{H}/^{13}\text{C}$ correlation spectra were obtained by ^{13}C observation

(“direct detection”), but, with improvements in spectrometer phase and frequency stability, these are now almost always obtained by ^1H observation (“indirect detection,” an historical but misleading term). One common feature of heteronuclear correlation sequences is a pair of 90° pulses (^{13}C pulses for ^1H -detected sequences) at the beginning and end of the evolution period. By monitoring the evolution of magnetization during t_1 , the second 90° pulse acts as the equivalent of a phase-sensitive detector, allowing determination of frequency information for that nucleus, even though the receiver is tuned to detect the acquisition nucleus.

14.3.3 Absolute-Value Versus Phase-Sensitive Spectra

As mentioned in Sect. 14.2.4, a pulse sequence produces both real (absorption) and imaginary (dispersion) spectra. With certain pulse sequences (COSY being the most common), it is impossible to simultaneously have all peaks with the same phase since the sequence produces a mixture of absorption and dispersion peaks. In this case, an absolute-value (or magnitude-mode) spectrum is generated by squaring the real and imaginary spectra, summing them, and then taking the square root of the sum. This arbitrarily produces an apparent absorption spectrum with all peaks with the same phase. However, it does so at some cost in resolution since the dispersion components have broad “tails,” which broaden the peaks.

Fortunately, most pulse sequences generate phase-sensitive (or pure-absorption) spectra, with better resolution. These can be generated in one of two ways. The first, which is usually the method of choice on Varian/Agilent spectrometers, is to acquire two sets of spectra with a phase shift of 90° between them. This is done alternately to avoid one set suffering more than the other from any degradation of resolution over time. They are then co-processed to produce the 2D spectrum. The second, usually used on Bruker spectrometers, is to acquire a single set of spectra with twice as many time increments, but having a 90° phase shift between each successive spectrum. Both approaches, which effectively mimic the approaches used for quadrature detection on Varian/Agilent and early Bruker spectrometers, yield very similar results in the same time.

One interesting exception is provided by the HMBC sequence. In its original form, it was designed for “mixed-mode” processing: absolute value along F_2 but phase sensitive along F_1 (118). This gave improvements in both resolution and sensitivity over a full absolute-value mode spectrum. Later versions of HMBC using gradients (see next section) produced straight absolute-value spectra. However, some of the recent gradient-selected sequences again allow mixed-mode processing with its associated advantages, including better ^{13}C resolution.

14.3.4 Phase Cycling Versus Gradient Selection

Either phase cycling or gradient selection (sometimes called gradient enhancement) or some combination of the two is used for two purposes in any multi-pulse NMR sequence. The first is coherence pathway selection and the second is artifact

suppression. The explanation of coherence pathway selection is well beyond the scope of this chapter but is discussed in various texts such as that by *Keeler (119)*. Instead we will focus on how the two techniques are carried out and their relative advantages and disadvantages. Phase cycling consists of varying the phase of one or more of the subsequent pulses and/or the receiver, relative to the phase of the initial pulse, in consecutive scans. The phase cycle is designed so that the desired signal co-adds in the different scans while other signals are cancelled. The phase of the initial pulse, designated x , is arbitrary, but the phases of subsequent pulses can be x , y , $-x$, or $-y$ (respectively corresponding to 0, 90, 180, or 270 degree phase shifts, relative to the first pulse). This is controlled by the timing circuit of the spectrometer. The receiver "phase cycling" is carried out by dividing the computer memory block in two sections and, respectively, adding the digitized signal to the first block, adding it to the second, subtracting it from the first or subtracting it from the second (corresponding to x , y , $-x$, and $-y$). The minimum possible phase cycle for coherence selection will be two scans, but incorporating artifact selection will require more scans (anywhere from 4 up to 16 or even higher).

In contrast, gradient selection uses a pair of magnetic field gradients applied along the z -axis. These are designed so that the magnetization associated with one coherence pathway is dephased by the first gradient but brought back in phase by the second for observation while other pathways are dephased by both gradients and not observed. Artifacts are eliminated in the same manner. There are two main advantages to gradient selection over phase cycling. First, it can often be carried out with as little as one scan, substantially decreasing the time needed to acquire a high sensitivity experiment such as COSY. Second, pathway selection and artifact suppression are carried out during each scan while phase cycling relies on subtraction of the data from one scan from another for artifact suppression. The latter is more susceptible to minor spectrometer instabilities. This is particularly important in ^1H -detected $^1\text{H}/^{13}\text{C}$ shift correlation spectra where one is detecting the 1.1% $^1\text{H}/^{13}\text{C}$ magnetization while suppressing the remaining ^1H magnetization. On the other hand, most gradient-selected sequences result in a $\sqrt{2}$ loss in overall sensitivity. Nevertheless, the other advantages of gradient selection are so great that most 2D experiments are now carried out with gradient selection. However, they often also incorporate some phase cycling.

14.3.5 Acquisition Times and Relaxation Delays

Two-dimensional spectra are generally carried out using much shorter acquisition times (*ca.* 0.1–0.4 s) than used for 1D spectra, both to save time and to keep the 2D data set to a reasonable size. However, this requires including a relaxation delay (*RD*) in order to allow at least partial recovery of z -magnetization before the next scan. In assessing a reasonable value for the relaxation delay, it is important to remember that, for ^1H -detected experiments in particular, relaxation will also be occurring during the acquisition time. Thus, the key parameter to optimize is the recycle time (*RT*), which is the sum of $AT + RD$. One can then afford to increase AT (and, therefore, increase F_2 resolution) by correspondingly decreasing RD . It has

been shown that the optimum compromise value of RT for most 2D sequences (other than NOESY and ROESY, see below) is *ca.* $1.3T_1$. Methyl groups typically have the longest relaxation times of all ^1H signals in a natural product, but they are usually by far the most intense signals. Thus, one can afford to sacrifice some intensity for methyl protons and choose a value of RT based on average T_1 values for methylene and methine protons (10).

The actual RT values to be used depend the molecular weight of the molecule since larger molecules have shorter relaxation times. For small natural products (200–350 Da), typical ^1H T_1 values are *ca.* 0.7–1.2 s, corresponding to RT s of 0.9–1.5 s. For molecules in the 350–500 Da range, we suggest RT values of 0.7–1.0 s and around 0.6 s for larger molecules. Assuming $AT \sim 0.2$ s, corresponding RD values would be 0.2 s less. These are all significantly less than those recommended in a book, which suggests parameters for a wide range of 2D experiments (8), but, as we have pointed out elsewhere (10), the authors regard the use of such long values of RT as a waste of spectrometer time.

14.3.6 Number of Time Increments, Forward Linear Prediction, and Zero Filling

As noted above, the acquisition time can be increased (and F_2 resolution improved) without increasing the total experiment time by correspondingly reducing the relaxation delay. In contrast, the total experiment time is directly proportional to the number of time increments used. Since many natural products have very crowded spectra, particularly in the aliphatic region, the authors find that one commonly needs 1,024 time increments (NI) to get satisfactory F_1 resolution. One can use a smaller number of increments in an attempt to save time, but the risk is that the spectrum will be too poorly resolved to allow unambiguous interpretation. Thus, rather than saving time, one actually has wasted it.

Fortunately, there is a well-established method that allows one to use a significantly smaller value of NI, thus saving time, while still getting adequate resolution. This is forward linear prediction (LP) (120). The idea behind LP can be likened to a race where different automobiles each travel at a different, but constant, speed. If their relative positions after 256 laps are noted, one can make a very good estimate of their relative positions after 512 or 1,024 laps. In NMR LP, finite-length interferograms are extended by using information from previous data points to predict additional data points. In a time sequence of data points, the value of a particular data point, $d(m)$, can be estimated from a linear combination (hence the name “linear prediction”) of the data points that immediately precede it (120):

$$d(m) = d(m-1)a(1) + d(m-2)a(2) + d(m-3)a(3) + \dots$$

where $a(1)$, $a(2)$, $a(3)$ are the LP coefficients. The number of coefficients used corresponds to the number of data points that are used to predict the value of the next data point in the series. When applying this method to phase-sensitive 2D data sets,

we have shown that one can reliably use fourfold LP, *e.g.* set NI=256 but linearly predict it out to NI=1024 (10, 121). This allows one to obtain comparable quality 2D spectra in one quarter of the time taken to obtain the full set of t_1 interferograms or to double the S/N in the same time by increasing NS by a factor of four. However, fourfold linear prediction does require a reasonable number of experimental time increments, with NI=128 the suggested lower limit. For smaller values of NI, only twofold LP is recommended. Also, the authors have found that only twofold LP could be reliably carried out for absolute-value (*e.g.* COSY) 2D spectra, even for NI=256 or larger (121).

The main requirement for the use of LP is that the latter parts of the experimental interferograms have sufficient S/N that an accurate estimation of the LP coefficients can be made. However, in over 20 years of processing well over one thousand 2D data sets with LP, we have rarely found this to be a problem. For example, we obtained an HSQC spectrum of a very dilute mixture of three polysaccharides using 16-fold linear prediction (NI=1024 out to NI=16384) and still obtained accurate data for very crowded spectra (122). For dilute solutions, due to relaxation, most of the signal intensity will occur in the first one quarter to one third of the interferogram with the signals in the later portions comparable to or even smaller than the random noise. Under these circumstances, the authors find that it is better to reduce NI (usually to $\frac{1}{4}$ of the desired value) while correspondingly increasing NS by a factor of four. In this way, S/N is improved (by a factor of 2) in the region of the interferogram that exceeds noise, improving the chances of successful LP of the rest of the interferogram. By contrast, if one omits LP and instead collects the full data set with the smaller value of NS, the entire interferogram may be too noisy to provide a good quality spectrum. If the former approach fails, it is probable that the sample is too dilute to obtain acceptable results in reasonable time by either method.

One additional requirement is that the number of LP coefficients should be greater than the number of signals that make up each interferogram that is being extended. The number of signals varies with the type of sequence, being as low as one or two with HSQC but generally larger. How much larger the number of coefficients should be for best quality spectra is spectrometer dependent. In our experience with Varian/Agilent spectrometers, LP works best when the number of LP coefficients is no larger than twice the expected number of signals while Bruker recommends that it should be at least two to three times larger. Too small a value may yield poor quality spectra and/or missing peaks while too large risks detecting spurious signals.

While LP is very useful for extending the number of t_1 time increments, it is of little value for extending NP in t_2 . First, particularly for ^1H -detected sequences, there are typically a very large number of F_2 signals, and thus LP would require a very large number of coefficients, which would seriously slow the calculations. Also, as noted above, one can increase NP (and AT) without any increase in experiment time by correspondingly reducing the relaxation delay. However, backward linear prediction can be used along F_2 . This can be used to correct any corrupted data points or to remove very broad background signals.

Finally, it should be realized that zero filling is not an alternative to LP but rather the two techniques are complementary. We recommend always using a further

equal amount of zero filling to the value of NI after LP (this is a requirement with Varian/Agilent LP software). This will further improve $F1$ data point resolution on phase-sensitive spectra by a factor of two. On the other hand, if one only used zero filling up to a factor of eight, the data point resolution would be still be four times worse than that if fourfold linear prediction were used in combination with twofold zero filling. In addition, one would have to use a more extreme apodization function to avoid artifacts due to truncation (see Sect. 14.3.8).

14.3.7 Number of Scans

The number of scans required for an acceptable spectrum depends on the sample concentration, the probe sensitivity and the type of sequence (including any essential phase-cycling requirements included in the sequence). For reasonable concentrations, gradient-selected COSY spectra can be obtained with $NS=1$, but most other sequences require a larger number of scans. This is particularly true for HMBC, which is the least sensitive of all of the pulse sequences used for organic structure elucidation.

14.3.8 Apodization Functions

Due to the short acquisition and evolution times used in 2D NMR, both FIDs and interferograms have typically not decayed away to zero at the end of t_2 or t_1 . Fourier transformation of a truncated FID or interferogram will result in a spectrum with distorted peaks due to “truncation wiggles.” To avoid this, it is essential to use an apodization (weighting) function that goes to zero at the end of the time period. The exact shape of the apodization function is mainly determined by whether the spectrum is obtained in absolute-value or phase-sensitive mode. To minimize the broad tails characteristic of absolute-value peaks, absolute-value spectra are processed by using a resolution-enhancement function, which starts at zero for t_1 or $t_2=0$, peaks at the middle, and again goes to zero at $t(\max)$. The most common forms are either a sine bell or a squared sine bell (or the near-equivalent pseudo-echo function). The squared sine bell gives slightly better resolution while the sine bell gives slightly better S/N . The former is recommended for COSY spectra and the latter for other, lower sensitivity, spectra. For phase-sensitive spectra, a function that starts at a maximum at $t=0$ and goes to zero at $t(\max)$ is recommended. Possibilities include a cosine (90° shifted sine bell) function, a *Gaussian* function or an exponential (line broadening) function. The first gives the best resolution, the last, the best S/N , while the *Gaussian* function is a compromise choice and is the one we generally prefer. A compromise function, which could be used for both absolute-value and phase-sensitive spectra, is a shifted sine bell function (typical shapes of this and other apodization functions are illustrated in Fig. 15). However, the authors find that it is less than ideal for either type of spectrum. In the case of an HMBC data set designed for mixed-mode processing, we recommend a sine bell along t_2 and a *Gaussian*

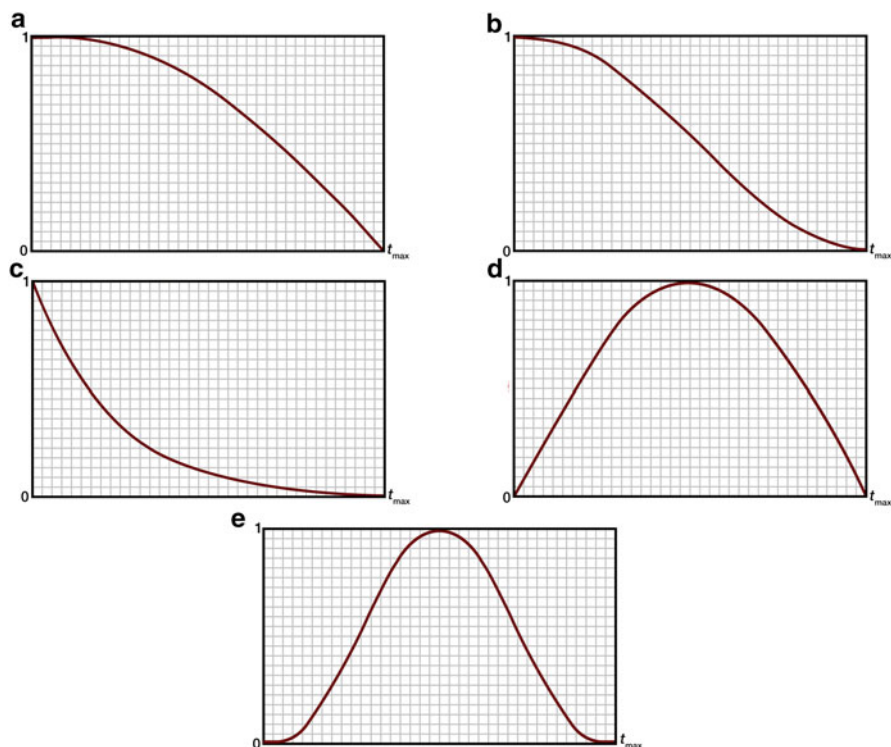


Fig. 15 The shapes of typical weighting functions used in processing 2D spectra. (a) cosine function 90° shifted sine bell function (b) *Gaussian* function (c) exponential (line broadening) function (d) sine bell function (e) squared sine bell function. Functions (a)–(c) are appropriate for phase-sensitive spectra while functions (d) and (e) are for absolute-value (magnitude-mode) spectra

along t_1 . Finally, if one is using linear prediction, it is important to remember that this effectively extends t_1 and that the chosen apodization function should be adjusted to be zero at the extended value of t_1 .

14.3.9 Data Point Resolution in 2D NMR Spectra

In both absolute-value and phase-sensitive 2D spectra, the F_2 data point resolution without zero filling is given by $2SW/NP$, identical to the value for 1D (see Sect. 14.2.4). An equal amount of zero filling of phase-sensitive spectra will again improve digital resolution to SW/NP . However, absolute-value spectra are different. Since both the real and imaginary points from the FID are used to generate an absolute-value spectrum, zero filling does not improve the digital resolution. For that reason, it is better to use a larger number of data points in F_2 for COSY and HMBC spectra in particular.

The situation for $F1$ data point resolution is more complex because it involves the way in which $F1$ quadrature detection is carried out. In the case of phase-sensitive spectra, the two alternative methods are closely analogous to the two methods for $F2$ quadrature detection described in Sect. 14.1.4. In the “States” method, two data sets are acquired, each with NI increments, with a 90° phase shift between them. FT yields two phase-sensitive spectra, which include both real and imaginary signals, but which can be phased to produce pure absorption-mode peaks. The difference is that one is a “cosine” spectrum where peaks of the same phase (one the true peak and the other the quadrature image peak) are mirrored about the center of $SW1$, the $F1$ spectral window. The second is a “sine spectrum” in which peaks are again mirrored about the middle of $SW1$, but now the quadrature image peak is of opposite phase to the true peak. When the two spectra are added, the quadrature image peaks cancel while the real peaks add. This co-addition of the two spectra improves the S/N by $\sqrt{2}$ but does not change the data point resolution, which is given by $2SW1/NI$ (since there are only $NI/2$ real peaks in each spectrum) or $SW1/NI$ with an equal amount of zero filling. In the case of the TPPI method, only a single data set is collected, but NI is doubled, with every second FID phase shifted by 90° . Although the actual processing method is different, the end result is the same. Effectively, one has collected the equivalent of two data sets of $NI/2$ increments. Thus, the digital resolution in this case is $4SW1/NI$ without zero filling or $2SW1/NI$ with zero filling. Allowing for the fact that NI is twice as large in the TPPI method, the actual data point resolution is identical.

In the case of absolute-value spectra, the data point resolution is again $2SW1/NI$ or $SW1/NI$ with one level of zero filling. However, in this case, $F1$ quadrature detection is carried out by using either phase cycling or gradient selection. This involves phase modulation of signals rather than amplitude modulation and produces signals with complex phase-twisted shapes. This is the reason that an absolute-value display is required. The need for one level of zero filling arises because carrying out quadrature detection with only a single data set results in only half of the points being used to generate the spectrum.

14.3.10 Shaped Pulses and Selective 1D Analogues of 2D NMR Spectra

Modern NMR spectrometers are equipped with wave form generators. These can be used to generate frequency-selective shaped RF pulses, including both 90° and 180° pulses (123). These are designed to generate uniform excitation over a defined spectral window, with ideally no excitation outside of this window. These are complex to design and difficult to implement manually. However, the software associated with spectrometer pulse sequence libraries usually makes this quite simple in practice, often using just two cursors to define the region to be excited, with the software then calculating the appropriate pulse shape.

Often in natural product research, one needs only correlation data for a limited number of protons to complete structural and/or stereochemical assignments. In these cases, using selective pulses to generate a series of 1D analogues of 2D

spectra may provide a considerable time saving. The most commonly used selective 1D sequences are 1D NOESY, ROESY, and TOCSY. One-dimensional TOCSY is particularly valuable since, by performing a series of measurements with increasing mixing times, one can sequentially trace out a network of coupled protons and potentially determining their coupling constants (*10*) (see Sect. 7 for an example), even when some of the protons overlap with other proton signals.

14.4 Recommended Acquisition and Processing Parameters for Commonly Used 2D Experiments and Selective 1D Experiments

The parameters listed below are designed to be appropriate for spectrometers in the 400–600 MHz range and equipped with either an indirect-detection probe or one of the newer probes with excellent sensitivity on both channels (Agilent OneNMR probe or Bruker SMART probe). If using an older direct-detection (¹³C-optimized) probe, an increased number of scans may be necessary for dilute solutions, while smaller numbers of scans are needed if using a cryogenically cooled probe. In each case, ranges of values for two key parameters are given: the number of scans (NS) and the relaxation delay (*RD*). In the case of NS, the minimum value is recommended when one has >5 mg of sample while the maximum is for cases with *ca.* 1 mg of sample. The minimum value of *RD* is suggested for compounds of molecular weight >750 Da while the maximum is for compounds of less than 300 Da. The recommended number of steady-state (dummy) scans is defined by SS prior to data acquisition. In addition, two different sets of recommendations are provided for the number of *F2* points (NP) and time increments (NI), the extent of linear prediction (LP) and the minimum extent of zero filling (ZF), with the latter two, respectively, defined as the total number of points after *F1* linear prediction and after zero filling. These are labeled “low resolution” and “high resolution” and are, respectively, suitable for compounds with clearly resolved ¹H spectra and spectra with one or more regions with significant spectral crowding. Finally, the parameters for the phase-sensitive spectra are those that are appropriate for acquisition using the “States” method. This is the standard choice on Varian/Agilent spectrometers and one of the options on Bruker spectrometers.

14.4.1 COSY and TOCSY Experiments

14.4.1.1 Gradient-Selected COSY (Absolute-Value Mode)

F1 = *F2* = –0.5 to 9.5 ppm, relative to TMS

SS = 16

NS = 1–4*

RD = 0.5–1.0 s

$F2$ apodization: sine bell squared

$F1$ apodization: sine bell squared

Low-resolution spectra: NP=1024, ZF(F2)=1024, NI=256, LP=512, ZF(F1)=1024

High-resolution spectra: NP=2048, ZF(F2)=2048, NI=512, LP=1024, ZF(F1)=2048

*If using the non-gradient (phase-cycled) version of COSY, NS must be a multiple of 4.

14.4.1.2 Gradient-Selected Double Quantum Filtered COSY (Phase Sensitive)

$F1 = F2 = -0.5$ to 9.5 ppm

SS=16

NS=1–8*

RD=0.5–1.0 s

$F2$ apodization: cosine (90° shifted sine bell)

$F1$ apodization: cosine (90° shifted sine bell)

High-resolution spectra**: NP=4096, ZF(F2)=8192, NI=256, LP=1024, ZF(F1)=2048

*If using the non-gradient DQCOSY sequence, NS must be a multiple of 4.

**Acquiring a high-resolution spectrum along $F2$ is strongly recommended since the main value of the experiment is its ability to measure coupling constants.

14.4.1.3 TOCSY or Z-TOCSY* (Phase-Sensitive)

$F1 = F2 = -0.5$ to 9.5 ppm

SS=16

NS=2–16

RD=0.5–1.0 s

Mixing time: 80 ms**

$F2$ apodization: *Gaussian*

$F1$ apodization: *Gaussian*

Low-resolution spectra: NP=1024, ZF(F2)=2048, NI=256, LP=1024, ZF(F1)=2048

High-resolution spectra: NP=2048, ZF(F2)=4096, NI=256, LP=1024, ZF(F1)=2048

*If available, the use of Z-TOCSY is recommended since the zero-quantum filter gives undistorted cross-peak patterns.

**A shorter mixing time (25–30 ms) will give a COSY-like spectrum but with better resolution than an absolute-value COSY spectrum.

14.4.2 NOESY and ROESY Experiments

14.4.2.1 NOESY (Phase-Sensitive)

$F1 = F2 = -0.5$ to 9.5 ppm

SS=16

NS=4–16

RD=1.0–1.8 s

Mixing time=0.3–0.8 s*

F2 apodization=*Gaussian*

F1 apodization=*Gaussian*

Low-resolution spectra: NP=1024, ZF(F2)=2048, NI=256, LP=1024, ZF(F1)=2048

High-resolution spectra: NP=2048, ZF(F2)=4096, NI=256, LP=1024, ZF(F1)=2048

*The choice of mixing time is critical. Small molecules need long mixing times to allow for reasonable NOE build-up. However, with high (>750 Da) molecules, spin-diffusion will occur with long mixing times, leading to erroneous results.

14.4.2.2 ROESY *(Phase-Sensitive)

F1 = F2 = -0.5 to 9.5 ppm

SS=16

NS=4–16

RD=1.0–1.8 s

Mixing time=0.2–0.6 s

F2 apodization=*Gaussian*

F1 apodization=*Gaussian*

Low-resolution spectra: NP=1024, ZF(F2)=2048, NI=256, LP=1024, ZF(F1)=2048

High-resolution spectra: NP=2048, ZF(F2)=4096, NI=256, LP=1024, ZF(F1)=2048

*The use of ROESY in place of NOESY is strongly recommended for molecules with molecular weights >600 Da since NOESY cross-peaks approach zero intensity for molecules much above this molecular weight, eventually become negative as molecular size increases.

14.4.3 HMQC, HSQC, HMBC, and H2BC Experiments

14.4.3.1 Gradient-Selected HMQC (Absolute-Value)

F2 = -0.5 to 9.5 ppm

F1 = -5 to 165 ppm

SS=16

NS=4–32

RD=0.5–1.0 s

$^1J_{\text{CH}}$ = 145 Hz

F2 apodization=sine bell

F1 apodization=sine bell

Low-resolution spectra: NP=1024, ZF(F2)=2048, NI=256, LP=512, ZF(F1)=1024

High-resolution spectra: NP=2048, ZF(F2)=4096, NI=512, LP=1024, ZF(F1)=2048

14.4.3.2 Gradient-Selected HSQC* (With or Without ^{13}C Spectral Editing) $F2 = -0.5$ to 9.5 ppm $F1 = -5$ to 165 ppm

SS = 16

NS = 4–32

RD = 0.5–1.0 s

 $^1J_{\text{CH}} = 145$ Hz $F2$ apodization = *Gaussian* $F1$ apodization = *Gaussian*

Low-resolution spectra: NP = 1024, ZF(F2) = 2048, NI = 160, LP = 512, ZF(F1) = 1024

High-resolution spectra: NP = 2048, ZF(F2) = 4096, NI = 256, LP = 1024, ZF(F1) = 2048

*HSQC gives better resolution than HMQC and allows spectral editing. Therefore, the use of HSQC in place of HMQC is strongly recommended.

14.4.3.3 Gradient-Selected HMBC (Absolute-Value)

 $F2 = -0.5$ to 9.5 ppm $F1 = -5$ to 220 ppm*

SS = 16

NS = 16–64

RD = 0.5–1.0 s

 $^1J_{\text{CH}} = 145$ Hz (or 130 Hz, 165 Hz)** $^nJ_{\text{CH}} = 8$ Hz $F2$ apodization = sine bell $F1$ apodization = sine bell

Low-resolution spectra: NP = 2048***, ZF(F2) = 4096, NI = 256, LP = 512, ZF(F1) = 1024

High-resolution spectra: NP = 4096, ZF(F2) = 8192, NI = 512, LP = 1024, ZF(F1) = 2048

*A high-frequency value of 200 ppm can be substituted if one is certain that the compound has no carbonyl groups.

**The values in parentheses are lower and upper values if a two-step J -filter is used.

***The use of at least a 0.2-s acquisition time is essential to avoid significant sensitivity loss, see (12).

14.4.3.4 Gradient-Selected HMBC (Mixed-Mode Processing)*

 $F2 = -0.5$ to 9.5 ppm $F1 = -5$ to 220 ppm**

SS = 16

NS = 16–64

RD = 0.5–1.0 s

 $^1J_{\text{CH}} = 145$ Hz (or 130 Hz, 165 Hz)*** $^nJ_{\text{CH}} = 8$ Hz

F2 apodization = sine bell

F1 apodization = *Gaussian*

Low-resolution spectra: NP = 2048, ZF(F2) = 4096, NI = 160, LP = 512, ZF(F1) = 1024

High-resolution spectra: NP = 4096, ZF(F2) = 8192, NI = 256, LP = 1024, ZF(F1) = 2048

*This sequence gives better *F1* resolution and better sensitivity than the absolute-value sequence. The spectra are in displayed in absolute-value mode but are processed in phase-sensitive mode along *F1*.

**A high frequency value of 200 ppm can be substituted if one is certain that the compound has no carbonyl groups.

***The values in parentheses are lower and upper values if a two-step *J*-filter is used.

14.4.3.5 Gradient-Selected H2BC (Phase-Sensitive)

F2 = -0.5 to 9.5 ppm

F1 = -5 to 220 ppm

SS = 16

NS = 16-64

RD = 0.5-1.0 s

$^1J_{\text{CH}}$ = 145 Hz (or 130 Hz, 165 Hz)

T(fixed time) = 0.022 s

F2 apodization = *Gaussian*

F1 apodization = *Gaussian*

Low-resolution spectra: NP = 1024, ZF(F2) = 2048, NI = 160, LP = 512, ZF(F1) = 1024

High-resolution spectra: NP = 2048, ZF(F2) = 4096, NI = 256, LP = 1024, ZF(F1) = 2048

14.4.4 Selective 1D Experiments

14.4.4.1 1D TOCSY*

F2 = -0.5 to 9.5 ppm

NP = 32,768

FN = 65,536

SS = 16

NS = 4-64**

Mix = 0.00 s and 0.08 s (or array)***

Apodization = 0.5-Hz line broadening

*If available, the use of the Z-TOCSY sequence is strongly recommended.

**If using a relatively long mixing time, a larger number of scans will be needed since the initial magnetization is spread amongst several proton multiplets.

***Acquisition of an initial spectrum with zero mixing time is recommended to ensure that one has a clean excitation of only the desired multiplet. Arraying the mixing time (e.g. 0.0, 0.25, 0.5, 0.75, 1.0 s) is useful since this allows one to assign sequences of coupled protons.

14.4.4.2 1D NOESY or ROESY*

$F_2 = -0.5$ to 9.5 ppm

NP = 32,768

FN = 65,536

SS = 16

NS = 16–256**

Mix = 0.5 s

Apodization = 2-Hz line broadening

*The use of NOESY is suggested for molecules of molecular weight <600 Da while ROESY is strongly recommended for larger molecules.

**Both NOESY and ROESY measure transient NOE buildup and are relatively insensitive.

15 Conclusions

As indicated in the Introduction, NMR spectroscopy is a very powerful tool for natural product structure elucidation. However, in order to obtain the best results in the shortest time and to avoid making errors in structure determination, it is important to make informed choices of pulse sequences, acquisition parameters, and processing parameters. It is also important to approach unknown structural problems with an open mind and let the data point you to the correct structure rather than trying to force the data to fit a structure that you suspect. Hopefully, this contribution will provide a natural product chemist, who has a basic understanding of NMR spectroscopy, with the increased knowledge needed to apply this technique more effectively in his/her research. If so, we will have achieved our goal in writing this chapter.

Acknowledgment The authors are indebted to *Arvin Moser* of Advanced Chemistry Development, Inc. for the expert analysis of 1- and 2-dimensional NMR spectral data of three natural products in the ACD/Labs *Structure Elucidator* program.

References

1. Openshaw HT, Robinson R (1946) Constitution of Strychnine and the Biogenetic Relationship of Strychnine and Quinine. *Nature* 157:438
2. Woodward RB, Cava MP, Ollis WD, Hunger A, Daeniker HU, Schenker K (1954) The Total Synthesis of Strychnine. *J Am Chem Soc* 76:4749
3. Ernst RR, Anderson WA (1966) Application of *Fourier* Transform Spectroscopy to Magnetic Resonance. *Rev Sci Instrum* 37:93
4. Aue WP, Bartholdi E, Ernst RR (1976) Two-Dimensional Spectroscopy. Application to Nuclear Magnetic Resonance. *J Chem Phys* 64:2229

5. Ernst RR (1992) Nuclear Magnetic Resonance *Fourier* Transform Spectroscopy (*Nobel Prize Lecture*). *Angew Chem Int Ed* 31:805
6. Silverstein RM, Webster FX, Kiemle D (2005) *Spectrometric Identification of Organic Compounds*, 7th ed. John Wiley and Sons, Inc., New York
7. Lambert JB, Gronert S, Shurvell HF, Lightner DA, (2011) *Organic Structural Spectroscopy*, 2nd ed. Prentice Hall, Upper Saddle River, NJ, USA
8. Berger S, Braun S (2004) *200 and More NMR Experiments. A Practical Course*. Wiley-VCH, Chichester, UK
9. Lambert JB and Mazzola EP (2004) *Nuclear Magnetic Resonance Spectroscopy. An Introduction to Principles, Applications and Experimental Methods*, Chapter 7. Pearson Education Inc., Upper Saddle River, NJ, USA
10. Reynolds WF, Enriquez RG (2002) Choosing the Best Pulse Sequences, Acquisition Parameters, Postacquisition Processing Strategies, and Probes for Natural Product Structure Elucidation by NMR Spectroscopy. *J Nat Prod* 65:221
11. Kwan EE, Huang SG (2008) Structural Elucidation with NMR Spectroscopy: Practical Strategies for Organic Chemists. *Eur J Org Chem*:2671
12. Burrow TE, Enriquez RG, Reynolds WF (2009) The Signal/Noise of an HMBC Spectrum Can Depend Dramatically Upon the Choice of Acquisition and Processing Parameters. *Magn Reson Chem* 47:1086
13. McLean S, Perpich-Dumont M, Reynolds WF, Jacobs H, Lachmarsing SS (1987) Unambiguous Structural and Nuclear Magnetic Resonance Spectral Characterization of Two Triterpenoids of *Maprounea guianensis*. *Can J Chem* 65:2519
14. NMRwiki.org/wiki/index.php?title=Databases
15. Barna JCI, Laue ED, Mayger MR, Skilling J, Worrall SJP (1987) Exponential Sampling, an Alternative Method for Sampling in Two-Dimensional NMR Experiments. *J Magn Reson* 73:69
16. Gray G (2011) Non-Uniform Sampling and CLEAN Processing in VnmrJ 3.2, Spinsights, Agilent Technologies, Inc., Santa Clara, CA, USA
17. Schanda P, Kupce E, Brutscher B (2005) SOFAST-HMQC Experiments for Recording Two-Dimensional Heteronuclear Correlation Spectra of Proteins Within a Few Seconds. *J Biomol NMR* 33:199
18. Kupce E, Freeman R (2007) Fast Multidimensional NMR by Polarization Sharing. *Magn Reson Chem* 45:2
19. Farias J (2007) ASAP-HMQC – Quick Multidimensional Spectra. Agilent Application Note 5990-9237EN, Agilent Technologies, Inc., Santa Clara, CA, USA
20. Pauli GF, Jaki BU, Lankin DC (2005) Quantitative ¹H NMR: Development and Potential of a Method for Natural Products Analysis. *J Nat Prod* 68:133
21. Pauli GF, Jaki BU, Lankin DC (2007) A Routine Experimental Protocol for qHNMR Illustrated With Taxol. *J Nat Prod* 70:589
22. www.qnmr.org
23. Jaki BU, Franzblau SG, Chadwick LR, Lankin DC, Zhang F, Wang Y, Pauli GF (2008) Purity-Activity Relationships of Natural Products: The Case of the Anti-TB Active Ursolic Acid. *J Nat Prod* 71:1742
24. Akoka S, Barantin L, Trierweiler M (1999) Concentration Measurement by Proton NMR Using the ERETIC Method. *Anal Chem* 71:2554
25. Wider G, Dreier L (2006) Measuring Protein Concentrations by NMR Spectroscopy. *J Am Chem Soc* 128:2571
26. Farrant RD, Hollerton JC, Lynn SM, Provera S, Sidebottom PJ, Upton RJ (2010) NMR Quantification Using an Artificial Signal. *Magn Reson Chem* 48:753
27. Crouch R, Russell D (2011) Easy, Precise and Accurate Quantitative NMR. Agilent Application Note 5990-7601EN, Agilent Technologies, Inc., Santa Clara, CA, USA
28. Shoolery JN (1984) Recent Developments in ¹³C- and Proton NMR. *J Nat Prod* 47:226
29. Wynants C, Hallenga K, Van Binst G, Michel A, Zanin J (1984) Assignment of Amino Acids in Proteins by Correlation of α -Hydrogen and Carbonyl Carbon-13 Resonances. *J Magn Reson* 57:93

30. Kessler H, Griesenger C, Zarbock J, Loosli HR (1984) Assignment of Carbonyl Carbons and Sequence Analysis in Peptides by Heteronuclear Shift Correlation *via* Small Coupling Constants with Broadband Decoupling in *t*1 (COLOC). *J Magn Reson* 57:331
31. Reynolds WF, Enriquez RG, Escobar LI, Lozoya X (1984) Total Assignment of ^1H and ^{13}C Spectra of Kauradien-9(11),16-oic Acid with the Aid of Heteronuclear Correlated 2D Spectra Optimized for Geminal and Vicinal ^{13}C - ^1H Coupling Constants: or What to Do When "INADEQUATE" is Impossible. *Can J Chem* 62:2421
32. Jacobs H, Ramdayal F, Reynolds WF, Mclean S (1986) Guyanin, a Novel Tetranortriterpenoid. Structure Elucidation by 2-D NMR Spectroscopy. *Tetrahedron Lett* 27:1453
33. Bax A, Freeman R (1981) Investigation of Complex Networks of Spin-Spin Coupling by Two-Dimensional NMR. *J Magn Reson* 44:542
34. Bax A, Davis DG (1985) MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy. *J Magn Reson* 65:355
- 35a. Bax A, Morris GA, (1981) An Improved Method for Heteronuclear Chemical Shift Correlation by Two-Dimensional NMR. *J Magn Reson* 42:501
- 35b. Bax A (1983) Broadband Homonuclear Decoupling in Heteronuclear Shift Correlation NMR Spectroscopy. *J Magn Reson* 53:517
- 36a. Reynolds WF, Hughes DW, Perpick-Dumont M, Enriquez RG (1985) A Pulse Sequence for Establishing Carbon-Carbon Connectivities *via* Indirect ^{13}C - ^1H Polarization Transfer Modulated by Vicinal ^1H - ^1H Coupling. *J Magn Reson* 63:413
- 36b. Reynolds, WF, McLean S, Perpick-Dumont M, Enriquez RG (1989) Improved ^{13}C - ^1H Shift Correlation Spectra for Indirectly Bonded Carbons and Hydrogens: the FLOCK Sequence. *Magn Reson Chem* 27:162
37. Bax A, Subramanian S (1986) Sensitivity-Enhanced Two-Dimensional Heteronuclear Shift Correlation NMR Spectroscopy. *J Magn Reson* 67:565
38. Bodenhausen G, Ruben DJ (1980) Natural Abundance Nitrogen-15 NMR by Enhanced Heteronuclear Spectroscopy. *Chem Phys Lett* 69:185
39. Bax A, Summers MF (1986) ^1H and ^{13}C Assignments From Sensitivity-Enhanced Detection of Heteronuclear Multiple-Bond Connectivity by 2D Multiple Quantum NMR. *J Am Chem Soc* 108:2093
40. Lambert JB, Mazzola EP (2004) Nuclear Magnetic Resonance Spectroscopy: An Introduction to Principles, Applications and Experimental Methods, Chapter 8. Pearson Education, Inc., Upper Saddle River, NJ, USA
41. Kock M, Kerssebaum, R, Bermel W (2003) A Broadband ADEQUATE Pulse Sequence Using Chirp Pulses. *Magn Reson Chem* 41:65
42. Meyer SW, Kock M (2008) NMR Studies of Phakellins and Isophakellins. *J Nat Prod* 71:1524
43. Cheatham SF, Kline M, Sasaki RR, Blinov KA, Elyashberg ME, Molodtsov SG (2010) Enhanced Automated Structure Elucidation by Inclusion of Two-Bond Specific Data. *Magn Reson Chem* 48:571
44. Martin GE, Hilton BD, Blinov KA (2011) HSQC-ADEQUATE Correlation: a New Paradigm for Establishing a Molecular Skeleton. *Magn Reson Chem* 49:248
- 45a. Zhang F, Bruschweiler R (2004) Indirect Covariance NMR Spectroscopy. *J Am Chem Soc* 126:13180
- 45b. Snyder DA, Bruschweiler R (2009) Generalized Indirect Covariance NMR Formulism for Establishment of Multidimensional Spin Correlations. *J Phys Chem A* 113:12898
46. Martin GE, Hadden CE (2000) Long-Range ^1H - ^{15}N Heteronuclear Shift Correlation at Natural Abundance. *J Nat Prod* 63:543
47. Martin GE, Williams AJ (2005) Long-Range ^1H - ^{15}N Heteronuclear Shift Correlation. *Ann Rep NMR Spectrosc* 55:1
48. Martin GE, Hilton BD, Moskau D, Freytag N, Kessler K, Colson K (2010) Long-Range ^1H - ^{15}N Heteronuclear Shift Correlation Across Wide F1 Spectral Windows. *Magn Reson Chem* 48:935
49. Kline M, Cheatham S (2003) A Robust Method for Determining ^1H - ^{15}N Long-Range Correlations: ^{15}N Optimized CIGAR-HMBC Experiments. *Magn Reson Chem* 41:307

50. Nicolaou KC and Snyder SA (2005) Chasing Molecules That Were Never There: Misassigned Natural Products and the Role of Chemical Synthesis in Modern Structure Elucidation. *Angew Chem Int Ed* 44:1012
- 51a. Schlegel B, Hartl A, Dahse H-M, Gollmick FA, Grafe U, Dorfelt H, Kappes B (2002) Hexacyclinol, a New Antiproliferative Metabolite of *Panus rudis* HKI 0254. *J Antibiot* 55:814
- 51b. La Clair JJ (2006) Total Synthesis of Hexacyclinol, 5-*epi*-Hexacyclinol, and Desoxocyclohexacyclinol Unveil an Antimalarial Motif. *Angew Chem Int Ed* 45:2769
- 51c. Rychnovsky SD (2006) Predicting NMR Spectra by Computational Methods: Structure Revision of Hexacyclinol. *Org Lett* 8:2895
- 51d. Porco, Jr JA, Su S, Lei X, Bardhan S, Rychnovsky SD (2006) Total Synthesis and Structure Assignment of (+)-Hexacyclinol. *Angew Chem Int Ed* 45:5790
52. Laszlo P (1967) Solvent Effects and Nuclear Magnetic Resonance. *Prog NMR Spectrosc* 3:231
53. Hansen PE (1981) Carbon-Hydrogen Spin-Spin Coupling Constants. *Prog NMR Spectrosc* 14:175
54. Tinto WF, Reynolds WF (unpublished results)
55. D'Armas HT, Mootoo BS, Reynolds WF (2000) An Unusual Sesquiterpene Derivative from the Caribbean Gorgonian, *Pseudopterogorgia rigida*. *J Nat Prod* 63:1593
56. Elyashberg ME, Blinov KA, Williams AJ, Molodtsov SG, Martin GE, Martirosian ER (2004) *Structure Elucidator*: A Versatile Expert System for Molecular Structure Elucidation from 1D and 2D NMR Data and Molecular Fragments. *J Chem Inf Comput Sci* 44:771
57. Blinov KA, Carlson D, Elyashberg, ME, Martin GE, Martirosian ER, Molodtsov S, Williams AJ (2003) Computer-Assisted Structure Elucidation of Natural Products with Limited 2D NMR Data: Application of the StructEluc System. *Magn Reson Chem* 41:359
58. Karplus M (1963) Vicinal Proton Coupling in Nuclear Magnetic Resonance. *J Am Chem Soc* 85:2870
59. Neuhaus D, Williamson MP (2000) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, 2nd Edition. Wiley, New York
60. Haasnoot CAG, DeLeeuw FAAM, Altona C (1980) The Relationship Between Proton-Proton NMR Coupling Constants and Substituent Electronegativities—I: An Empirical Generalization of the *Karplus* Equation. *Tetrahedron* 36:2783
61. Musher JJ, Corey, EJ (1962) Virtual Long-Range Spin-Spin Couplings in NMR: The Linear 3-Spin System and Qualitative Implications of Higher Systems. *Tetrahedron* 18:791
- 62a. Leon I, Enriquez RG, McLean S, Reynolds WF, Yu M (1998) Isolation and Identification by 2D NMR of Two New Complex Saponins from *Michroseechium helleri*. *Magn Reson Chem* 36:S111
- 62b. Powder-George Y, Frank J, Ramsewak RS, Reynolds WF (2012) The Use of Coupled HSQC Spectra to Aid in Stereochemical Assignment of Molecules with Severe Proton Spectral Overlap. *Phytochem Anal* 23:274
63. Mazzola EP, Parkinson A, Kennelly EJ, Coxon B, Einbond LS, Freedberg DI (2011) Utility of Coupled-HSQC Experiments in the Intact Structural Elucidation of Three Complex Saponins from *Blighia sapida*. *Carbohydr Res* 346:759
64. Turner JJ, Shephard N (1959) High-Resolution Nuclear-Magnetic-Resonance Spectra of Hydrocarbon Groupings. II. Internal Rotation in Substituted Ethanes and Cyclic Ethers. *Proc Roy Soc A* 252:506
65. Macura S, Huang Y, Suter D, Ernst RR (1981) Two-Dimensional Chemical Exchange and Cross-Relaxation Spectroscopy of Coupled Nuclear Spins. *J Magn Reson* 43:259
66. Bothner-By AA, Stephens RL, Lee J-M, Warren CD, Jeanloz RW (1984) Structure Determination of a Tetrasaccharide: Transient Nuclear *Overhauser* Effects in the Rotating Frame. *J Am Chem Soc* 106:811
- 67a. Stott K, Stonehouse J, Keeler J, Hwang T-L, Shaka AJ (1995) Excitation Sculpting in High-Resolution Nuclear Magnetic Resonance Spectroscopy: Application to Selective NOE Experiments. *J Am Chem Soc* 117:4199

- 67b. Thrippleton MJ, Keeler J (2003) Elimination of Zero-Quantum Interference in Two Dimensional NMR Spectra. *Angew Chem Int Ed* 42:3938
- 68a. Butts CP, Jones CR, Towers EC, Flynn JL, Appleby L, Barron NJ (2011) Interproton Distance Determinations by NOE – Surprising Accuracy and Precision in a Rigid Organic Molecule. *Org Biomol Chem*. 9:177
- 68b. Butts CP, Jones CR, Harvey JN (2011) High Precision NOEs as a Probe for Low Level Conformers – A Second Conformation of Strychnine. *Chem Commun* 47:1193
69. Macura S, Farmer BT, Brown LR (1986) An Improved Method for the Determination of Cross-Relaxation Rates from NOE Data. *J Magn Reson* 70:493
70. Hu H, Krishnamurthy K (2006) Revisiting the Initial Rate Approximation in Kinetic NOE Measurements. *J Magn Reson* 182:173
71. Kummerlowe G, Luy B (2009) Residual Dipolar Couplings for the Configurational and Conformational Analysis of Organic Molecules. *Ann Rep NMR Spectrosc* 68:193
72. Gil RR (2011) Constitutional, Configurational, and Conformational Analysis of Small Organic Molecules on the Basis of NMR Residual Dipolar Couplings. *Angew Chem Int Ed* 50:7222
73. Kummerlowe G, Crone B, Kretschmer M, Kirsch SF, Luy B (2011) Residual Dipolar Couplings as a Powerful Tool for Constitutional Analysis: The Unexpected Formation of Tricyclic Compounds. *Angew Chem Int Ed* 50:2643
74. Jones CR, Butts CP, Harvey JN (2011) Accuracy in Determining Interproton Distances Using Nuclear *Overhauser* Effect Data from a Flexible Molecule. *Beilstein J Org Chem* 7:145
75. Dale JA, Mosher HS (1973) Nuclear Magnetic Resonance Enantiomer Reagents. Configurational Correlations *via* Nuclear Magnetic Resonance Chemical Shifts of Diastereomeric Mandelate, *O*-Methylmandelate, and α -Methoxy- α -trifluoromethylphenylacetate (MTPA) Esters. *J Am Chem Soc* 95:512
76. Collins DO, Reynolds WF, Reese PB (2004) New Cembranes from *Cleome spinosa*. *J Nat Prod* 67:179
77. Molina-Salinas GM, Rivas-Galindo VM, Said-Fernandez S, Lankin DC, Munoz MA, Joseph-Nathan P, Pauli GF, Waksman N (2011) Stereochemical Analysis of Leubethanol, an Anti-TB-Active Serrulatane from *Leucophyllum frutescens*. *J Nat Prod* 74:1842
78. Piozzi F, Passannanti S, Marino ML, Spiro V (1972) Structure of Grandiflorenic Acid. *Can J Chem* 50:109
79. Lederberg J, Sutherland GL, Buchanan BG, Feigenbaum EA, Robertson AV, Duffield AM, Djerassi C (1969) Applications of Artificial Intelligence for Chemical Inference. I. The Number of Possible Organic Compounds. Acyclic Structures Containing C, H, O and N. *J Am Chem Soc* 91:2973
80. Sasaki S-I, Abe H, Ouki T, Sakamoto M, Ochiai S (1968) Automated Structure Elucidation of Several Kinds of Aliphatic and Alicyclic Compounds. *Anal Chem* 40:2220
81. Nelson DB, Munk ME, Gash KB, Herald DL (1969) Alanylactinobicyclone. An Application of Computer Techniques to Structure Elucidation. *J Org Chem* 34:3800
82. Elyashberg ME, Gribov LA (1968) Formal-Logical Model for Interpreting Infrared Spectra From Characteristic Frequencies. *J Appl Spectrosc* 8:189
83. Christie BD, Munk ME (1987) The Application of Two-Dimensional Nuclear Magnetic Resonance Spectroscopy in Computer-Assisted Structure Elucidation. *Anal Chim Acta* 200:347
84. Nuzillard J-M, Massiot G (1991) Logic for Structure Determination. *Tetrahedron* 47:3655
85. Peng C, Yuan S, Zheng C, Hui, Y (1994) Efficient Application of 2D Correlation Information in Computer-Assisted Structure Elucidation of Complex Natural Products. *J Chem Inf Comput Sci* 34:805
86. Steinbeck C (1996) LUCY – A Program for Structure Elucidation from NMR Correlation Experiments. *Angew Chem Int Ed* 35:1984
87. Lindel T, Junker J, Koeck M (1999) 2D-NMR-Guided Constitutional Analysis of Organic Compounds Employing the Computer Program COCON. *Eur J Org Chem* 3:573

88. Moser A, Elyashberg ME, Williams AJ, Blinov KA, DiMartino JC (2012) Blind Trials of Computer-Assisted Structure Elucidation Software. *J Cheminformatics* 4:5
89. McLean S, Perpich-Dumont M, Reynolds WF, Sawyer JF, Jacobs H, Ramdayal F (1988) Guyanin, a Novel Tetranortriterpenoid: Structural Characterization by 2D NMR Spectroscopy and X-ray Crystallography. *J Am Chem Soc* 110:5339
90. Bamburg JR, Riggs NV, Strong FM (1968) The Structures of Toxins from Two Strains of *Fusarium tricinctum*. *Tetrahedron* 24:3329
91. Liler M (1971) Studies of Nuclear Magnetic Resonance Chemical Shifts Caused by Protonation. Part II. Formamide and Some *N*-Alkyl and *N,N*-Dialkyl Derivatives. *J Chem Soc B*:334
92. McClelland RA, Reynolds WF (1974) ^{13}C Nuclear Magnetic Resonance Spectra of *N,N*-Dimethylformamide in Aqueous Acid Solution. Evidence for Predominant *O*-Protonation at All Acidities. *J Chem Soc Chem Commun*:824
93. Burns D, Reynolds WF, Buchanan G, Reese PB, Enriquez RG (2000) Assignment of ^1H and ^{13}C Spectra and Investigation of Hindered Side-Chain Rotation in Lupeol Derivatives. *Magn Reson Chem* 38:488
94. Christian OE, Henry GE, Jacobs H, McLean S, Reynolds WF (2001) Prenylated Benzophenone Derivatives from *Clusia havetiodes* var. *stenocarpa*. *J Nat Prod* 64:23
95. Reynolds WF, McLean S, Tay L-L, Yu M, Enriquez RG, Estwick DM, Pascoe KO (1997) Comparison of ^{13}C Resolution and Sensitivity of HSQC and HMQC Sequences and Application of HSQC-Based Sequences to the Total ^1H and ^{13}C Spectral Assignment of Clianosterol. *Magn Reson Chem* 35:455
96. Kupce E, Freeman R (1995) Adiabatic Pulses for Wideband Inversion and Broadband Decoupling. *J Magn Reson A* 115:273
97. Wilker W, Leibfritz D, Kerssebaum R, Bermel W (1993) Gradient Selection in Inverse Heteronuclear Correlation Spectroscopy. *Magn Reson Chem* 31:287
98. Boyer RD, Johnson R, Krishnamurthy K (2003) Compensation of Refocusing Inefficiency with Synchronized Inversion Sweep (CRISIS) in Multiplicity-Edited HSQC. *J Magn Reson* 165:253
99. Hu H, Krishnamurthy K (2008) Doubly Compensated Multiplicity-Edited HSQC Experiments Utilizing Broadband Inversion Pulses. *Magn Reson Chem* 46:683
100. Reynolds WF, McLean S, Jacobs H, Harding WW (1999) Assignment of ^1H and ^{13}C Spectra for Polyprenol-12, a Molecule With Severe ^1H and ^{13}C Spectral Crowding, With the Aid of High-Resolution, ^{13}C -Detected, ^{13}C - ^1H Shift Correlation Spectra. *Can J Chem* 77:1922
101. Schoefberger W, Schlagnitweit J, Müller N (2011) Recent Developments in Heteronuclear Multiple-Bond Correlation Experiments. *Ann Rep NMR Spectrosc* 72:1
102. Furrer J (2011) Recent Developments in HMBC Studies. *Ann Rep NMR Spectrosc* 74:293
103. Hadden CE (2005) Adiabatic Pulses in ^1H - ^{15}N Direct and Long-Range Heteronuclear Correlations. *Magn Reson Chem* 43:330
104. Nyberg NT, Duss JO, Sorensen OW (2005) Heteronuclear Two-Bond Correlation: Suppressing Heteronuclear Three-Bond or Higher NMR Correlations while Enhancing Two-Bond Correlations Even for Vanishing $^2J_{\text{CH}}$. *J Am Chem Soc* 127:6154
105. Wagner R, Berger S (1998) ACCORD-HMBC: a Superior Technique for Structure Elucidation. *Magn Reson Chem* 36:S44
106. Hadden CE, Martin GE, Krishnamurthy VV (2000) Constant Time Inverse-Detection Gradient Accordion Rescaled Heteronuclear Multiple Bond Correlation Spectroscopy: CIGAR-HMBC. *Magn Reson Chem* 38:143
107. Furrer J (2010) A Robust, Sensitive and Versatile HMBC Experiment for Rapid Structure Elucidation by NMR: IMPACT-HMBC. *Chem Commun* 46:3396
108. Hansen DF, Kay LE (2007) Improved Magnetization Alignment Schemes for Spin-Lock Relaxation Experiments. *J Biomol NMR* 37:245
109. Exarchou V, Krucker M, van Beek TA, Vervoort J, Gerotheranassis IP, Albert K (2005) LC-NMR Coupling Technology: Recent Advancements and Applications in Natural Products Analysis. *Magn Reson Chem* 43:681

110. Tang H, Xiao C, Wang Y (2009) Important Roles of the Hyphenated HPLC-DAD-MS-SPE-NMR Technique in Metabonomics. *Magn Reson Chem* 47:S157
111. Martin GE (2005) Small-Volume and High-Sensitivity NMR Probes. *Ann Rep NMR Spectrosc* 56:1
112. Brey WW, Edison AS, Nast RE, Rocca JR, Saha S, Withers RS (2006) Design, Construction, and Validation of a 1-mm Triple Resonance High-Temperature-Superconducting Probe for NMR. *J Magn Reson* 179:290
113. Dalisay DS, Rogers EW, Edison AS, Molinski TF (2009) Structure Elucidation at the Nanomole Scale. 1. Trisoxazole Macrolides and Thiazole-Containing Cyclic Peptides from the Nudibranch *Hexabranchius sanguineus*. *J Nat Prod* 72:732
- 114a. Schroeder FC, Gronquist M (2006) Extending the Scope of NMR Spectroscopy with Microcoil Probes. *Ang Chem Int Ed* 45:7122
- 114b. Lambert M, Wolfender J-L, Staerk D, Christensen S B, Hostettmann K, Jaroszewski JW (2007) Identification of Natural Products Using HPLC-SPE Combined with CapNMR. *Anal Chem* 79:727
115. Patt SL, Shoolery JN (1982) Attached Proton Test for Carbon-13 NMR. *J Magn Reson* 46:535
116. Doddrell, DM, Pegg DT, Bendall MR (1982) Distortionless Enhancement of NMR Signals by Polarization Transfer. *J Magn Reson* 48:323
117. Burger R, Bigler P (1998) DEPTQ: Distortionless Enhancement by Polarization Transfer Including the Detection of Quaternary Nuclei. *J Magn Reson* 135:529
118. Bax A, Marion D (1988) Improved Resolution and Sensitivity in ¹H-Detected Heteronuclear Multiple-Bond Correlation Spectroscopy. *J Magn Reson* 78:186
119. Keeler J (2010) *Understanding NMR Spectroscopy*, 2nd Edition. Wiley, Chichester, UK
120. Hoch JC, Stern AS (1996) *NMR Data Processing*. Wiley-Liss, New York
121. Reynolds WF, Yu M, Enriquez RG, Leon I (1997) Investigation of the Advantages and Limitations of Forward Linear Prediction for Processing 2D Data Sets. *Magn Reson Chem* 35:505
122. Reynolds WF, Enriquez RG (2003) The Advantages of Forward Linear Prediction Over Multiple Aliasing for Obtaining High-Resolution HSQC Spectra in Systems With Extreme Spectral Crowding. *Magn Reson Chem* 41:927
123. Freeman R (1998) Shaped Radiofrequency Pulses in High Resolution NMR. *Prog NMR Spectrosc* 32:59