14

Software

There are many software tools available to help with the acquisition, processing and interpretation of NMR data. Attempts have been made to automate the verification process and even perform full structural elucidations of unknown compounds. As you might guess from the complexity of the interpretation chapters, these software solutions are not foolproof! It remains to be seen whether they ever will be good enough but there have certainly been some major steps forward in all of these areas.

In this chapter, we will look at the different types of software but be warned that software development is quite dynamic and the landscape may be very different when you come to read this section!

14.1 Acquisition Software

You seldom have much choice about this software. When you buy a spectrometer you will get some software from the manufacturer. The big manufacturers are Bruker, Varian and JEOL. Their software is called: Topspin, VNMRJ and Delta respectively. These pieces of software are quite complex as they have to perform all the spectrometer control as well as processing and some simulation. That said, all manufacturers have improved their software to make it more user-friendly in recent times and it is not the challenging beast that it used to be.

14.2 Processing Software

As mentioned above, the manufacturers provide software to process the data. These pieces of software are designed to process data created on that manufacturer's instrument although they can process most data from other vendors (sometimes this is not as easy as it could be). In addition to manufacturers' software, there are also third party software suppliers who offer software capable of processing data from all makes of NMR spectrometer (Figure 14.1). At the time of writing, there are a number of these companies, the most well known of these are probably Advanced Chemistry Development (ACD/Labs; http://www.acdlabs.com) and Mestrelab Research (http://www.mestrec.com). ACD's product for processing is called ACD/Specmanager (there are modules that you can purchase for 1-D or 2-D processing).

Essential Practical NMR for Organic Chemistry S. A. Richards and J. C. Hollerton
 ISBN: 978-0-470-71092-0

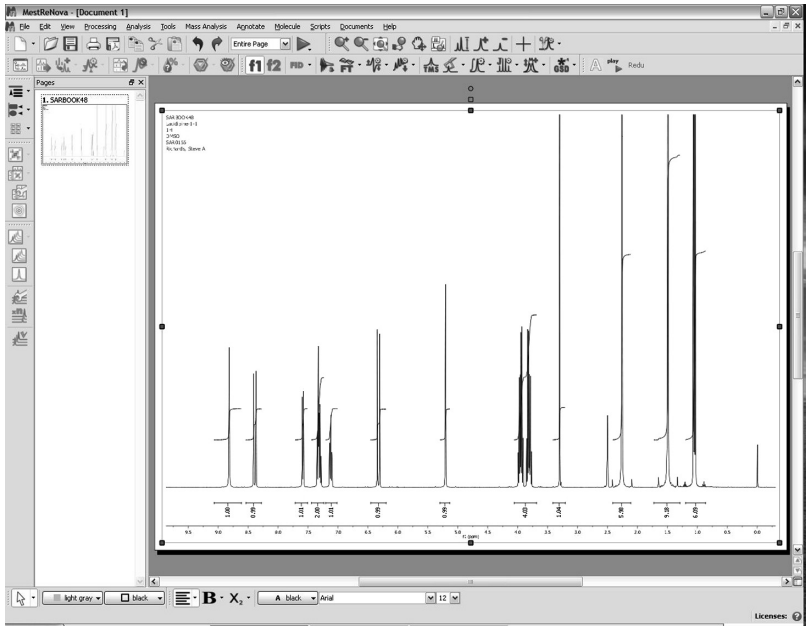

Figure 14.1 Third party processing software.

Mestrelab have a product called MestReNova which performs 1-D and 2-D processing. These products are focused on data processing and so tend to be a little less daunting than the spectrometer software.

There are many other pieces of processing software out there, some good, some bad. If you don't have any and you need some, it is worth having a look around to see what is available. In our experience, you tend to get what you pay for and the more expensive software is generally better. However, one of the cheaper packages may do just what you want, in which case you have a bargain! Table 14.1 lists some of the more well known packages.

Table 14.1 Processing software packages.

Name	Web site	Platform
NUTS	www.acornnmr.com	PC
NMRnotebook	www.nmrtec.com	Mac, PC, Linux
NMRPipe	spin.niddk.nih.gov	Unix, Linux
iNMR	www.inmr.net	Mac
MestreNova	www.mestrec.com	PC, Mac, Linux
ACD/SpecManager	www.acdlabs.com	PC
FELIX	www.felixnmr.com	PC, Linux, Unix

14.3 Prediction and Simulation Software

You will normally have access to the previous two categories of software as a minimum – they will be on the spectrometer itself. One thing that you don't get supplied is software to predict chemical shifts (although you may get some sort of simulation software). The desire to predict the chemical shift of a nucleus has been around since the first time that the chemical shift phenomenon was discovered. There are numerous papers going back to the earliest days of NMR trying to relate structural properties to chemical shift. Early work was concentrated on proton chemical shift prediction (because carbon data was so hard to get) but it was soon realized that the unpredictable nature of proton chemical shifts (their dependency on average solution conformation) made this job difficult. It was easy to get to within about 0.5 ppm of the correct shift but this is not too good when 90 % of your chemical shifts come within a 6 ppm range. Apart from generating additivity tables (as used in this book), proton chemical shift prediction was soon ignored.

14.3.1 ^{13}C Prediction

Once ^{13}C data was more readily available (with the advent of FT spectrometers), interest in chemical shift prediction was reborn. The reasons for this were that carbon spectra don't show carbon–carbon coupling information (unlike proton–proton coupling) and so knowledge of carbon chemical shifts was really important in the assignment of ^{13}C spectra. There were numerous efforts at carbon prediction but perhaps the first truly successful method was created by Wolfgang Bremser in 1977. He realized that you needed to have a standard way of naming and sorting carbon atoms so that you could look them up in a table. He also realized that if you were methodical about this and used a naming system that grew in 'shells' (Figure 14.2) from the atom of interest (atoms closest to the atom of interest come first in the name), you could predict the chemical shift of a similar compound by interpolating between entries in his table. He named this the HOSE code (*h*ierarchically *o*rdered *s*pherical description of *e*nvironment).

Bremser produced tables of these HOSE codes from NMR work that was carried out at BASF in Germany. Most modern carbon prediction routines still use this HOSE code today (albeit slightly modified from the original). Modern software hides all the HOSE code generation in the background so all you do is draw a structure and press the predict button and all the chemical shifts are calculated.

Modern carbon prediction software has hundreds of thousands of chemical structures to call on (Bremser had about 10 000 when he started). The more structures you have, the better the chance that something similar to your structure will be in the database - and the better the quality of the chemical shift

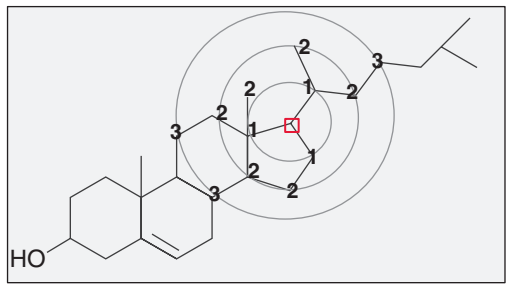

Figure 14.2 HOSE code in operation.

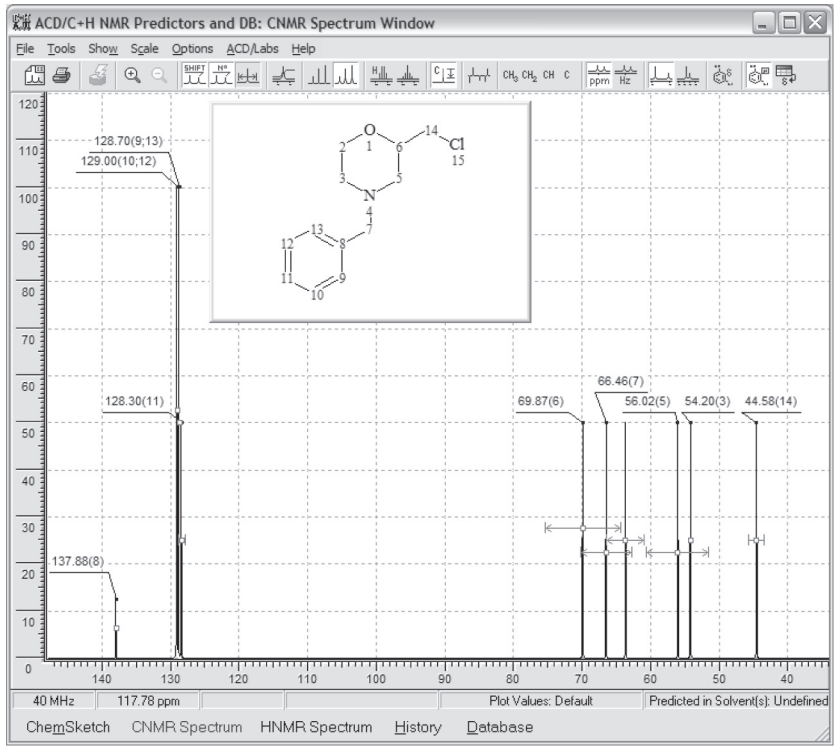

Figure 14.3 Carbon NMR prediction.

prediction. On the other hand, you may have something that has a fragment that just isn't represented in the database – in which case you cannot predict the chemical shift accurately.

There are a number of ^{13}C prediction packages that are commercially available. Once again, ACD/Labs offer a widely used prediction application called ACD/CNMR (Figure 14.3). This is well respected and has a large database with it. Another piece of software with a large database is NMRPredict from Modgraph. This software offers two different methods of ^{13}C prediction: a HOSE code method (with modification, designed by Wolfgang Robien) and a neural network approach (again from Wolfgang Robien). The HOSE code method works very well for well represented fragments but if you have an unusual fragment, the neural network offers better interpolation.

All in all, carbon prediction is really very good. This is partially due to the huge amount of carbon data in the public domain, partially due to the fact that carbon chemical shifts are distributed over 200 ppm (instead of 10 ppm for proton) and partially due to the fact that carbon chemical shift is mainly influenced by contributions through bond rather than through space (hence less dependence on conformation). Nonetheless, you need to be a little discriminating when assessing the chemical shift values that these systems come up with – if you have a carbon atom in an unusual environment and this is not covered in the database, the prediction bears very little weight and you must rely on other information to assign the atom.

14.3.2 ^{1}H Prediction

Despite having been the earliest attempted prediction, proton prediction remains relatively poor. The reasons for this have been alluded to earlier but to summarise; the proton chemical shift is often highly dependant on through-space effects (anisotropy) and has a very small distribution. There are four main commercial approaches to proton prediction currently: Incremental parameters, HOSE code databases, semi-empirical and *ab initio* methods.

14.3.2.1 Incremental Approaches

These are computerised versions of the tables in this book. Chemical shifts are calculated by adding together the contributions from the various functional groups attached to the core of interest. These are normally split into three types: aromatic, aliphatic and olefinic. Probably the best example of this approach is the Upsol predictor that was supplied with the book *Structure Determination of Organic Compounds* by Pretch *et al.* (Springer). This has found its way as an add-in to a few commercial systems and is currently used in the Modgraph NMRPredict package. The advantage of this approach is that calculations are quick and it is very easy to implement (and hence low cost). The disadvantage is that it is not very accurate and becomes progressively less accurate as more substituents are added to a core.

14.3.2.2 HOSE Code Databases

ACD/Labs have an extensive database which uses this approach. This approach works well except for anisotropic groups. Unlike carbon prediction this can have a massive effect on the chemical shift values and so can give rise to big errors in prediction, even for structural fragments that are well represented in the database.

14.3.2.3 Semi-Empirical Approaches

Currently there is only one product that adopts this approach and this is NMRPredict from Modgraph. It is based on the work by Prof. Ray Abraham at the University of Liverpool (UK). This approach calculates chemical shifts for a range of low energy conformers and averages them to give a net chemical shift. This approach seems to offer the most accurate prediction of chemical shift but the disadvantage is that it is very slow (particularly for conformationally flexible molecules).

14.3.2.4 **Ab Initio** *Approaches*

Another way to calculate chemical shifts is to use density functional theory (DFT). This quantum mechanical approach has been shown to predict chemical shifts well in certain cases. The disadvantage with this and semi-empirical approaches is that they rely on modelling the range of low energy conformers of the structure of interest. Not only is this time-consuming, it is also difficult to achieve in conformationally flexible molecules. Due to its slow performance, it is not a tool in regular use in solving structural problems although it has shown its value in specific cases, particularly where databases are not available for the structural feature/nucleus of interest.

14.3.3 Simulation

Spectral simulation is normally provided with proton prediction packages. This takes the predicted chemical shifts and coupling constants and uses them to simulate the appearance of the spectrum. This can be a little misleading as it gives rise to an authentic looking spectrum which may differ considerably in appearance to the experimental one. This is because even small errors in chemical shift or coupling constant prediction may give rise to significant differences in appearance of the signals. Simulation can be useful to try to mimic an observed signal to help calculate the coupling constants when they are not obvious by inspection. Simulation-only software is normally available as part of the NMR acquisition software and may be used to help understand complex splitting patterns observed in real spectra.

14.3.4 Structural Verification Software

Being able to verify a proposed chemical structure from its NMR spectrum automatically has been a goal for many years. This is particularly true recently since chemists have been making arrays of compounds (tens to thousands of compounds). It is possible to acquire data automatically on large numbers of compounds but it is still a major task to interpret all of the data. Verification software performs a prediction and simulation and then tries to fit the experimental data to the calculated data. Obviously this approach requires good prediction as well as good data extraction. As you will have seen in this book, these things are neither trivial nor reliable. The latest approaches use a combination of 1-D proton spectra and 2-D proton–carbon correlated spectra to try to use the strengths of ^{13}C prediction to aid the process. There appears to be some promise with this approach but it still has a way to go before it is truly reliable. The leaders in this area are currently ACD/Labs with a product called 'NMR Expert,' although a number of other companies are getting involved in this.

14.3.5 Structural Elucidation Software

Unlike structural verification software, this software is designed to propose structures that may fit the analytical data. The first requirement is for a molecular formula for the mystery compound. The more data that you have, the better, so you would normally need a proton–carbon correlated spectrum, plus any information that you can glean from the data (so if you can spot an ethyl group, you enter this information into the programme). You build up so-called 'good lists' and 'bad lists' of fragments that you might think are present and the software then calculates all the possible structures that could fit the data that you have entered. It then orders them by similarity between the experimental and calculated spectral data. This approach has been very useful in the area of natural product structural elucidation. Another area where this approach has potential is in spotting other possible structures that may fit the data. Even the best spectroscopists sometime become fixated with a structure that fits the data. This software can suggest other possible structures that are worthy of consideration.