

A beginner's guide to nuclear magnetic resonance: from atomic spies to complex 3D structures at the heart of structural biology

Timothy J. Woodman
(University of Bath, UK)

The 'Beginner's Guides' are an ongoing series of feature articles, each one covering a key technique and offering the scientifically literate, but not necessarily expert audience, a background briefing on the underlying science of a technique that is (or will be) widely used in molecular bioscience. The series will cover a mixture of techniques, including some that are well established amongst a subset of our readership but not necessarily familiar to those in different specialisms. This 'Beginner's Guide' covers nuclear magnetic resonance. Nuclear magnetic resonance (NMR) provides a way to explore the 3D structures of macromolecules in much more biologically relevant conditions and does this by taking advantage of the quantum mechanical property of some nuclei—nuclear spin. Here, we discuss how nuclear spin can be harnessed to provide information on the 3D structure of macromolecules in solution and how new thinking is leading to a revolution in drug discovery.

At the heart of NMR is the concept of nuclear spin, a quantum mechanical form of angular momentum. Nuclear spin has a direction and a magnitude, and was originally conceived as the rotation of a particle around an axis. However, this simplistic idea is complicated by the fact that the spin quantum number can take half integer numbers and that although the direction of the spin may be changed, the *rate* of the spin cannot be made faster or slower. In the presence of an external magnetic field, the simplest nucleus, a proton, tends to align either with or against the field. This leads to small energy differences between the two states, and typically this energy difference can be probed by choosing electromagnetic radiation of the 'right' frequency. For most modern systems the magnetic fields are roughly five orders of magnitude stronger than the Earth's magnetic field, and are only achieved through the use of superconducting materials cooled to near absolute

zero with liquid helium. NMR systems tend not to be described by the strength of the magnet, rather the frequency of electromagnetic radiation that is absorbed by hydrogen nuclei (protons) is chosen, and this varies from around 300 megahertz (MHz) up to 1.02 gigahertz (GHz). Unlike other forms of electromagnetic spectroscopy, the NMR effect is weak, as the signal depends on the tiny overpopulation of nuclei aligned with the field; for a 400 MHz magnet this is only a few nuclei per million. This intrinsic weakness requires the NMR signal produced by the nuclei to be extensively amplified prior to detection.

Most hydrogen nuclei in the universe are protons, and so are NMR active. The true utility of NMR is that it tells us not just that the proton is present in the sample, but also that the signal obtained depends on the local environment of the proton under study. The first piece of information regarding the local environment is termed

the chemical shift, and this captures small changes in the local magnetic environment of the nucleus arising from the way that other atoms are joined in chemical bonds. Electrons in a magnetic field also have spin and generate small magnetic fields as a result. These small fields can either reduce or enhance the external magnetic field, and as the frequency that the nucleus responds to (or 'resonates') is directly proportional to the magnetic field, so changes in the local magnetic field affect where the nuclei come into resonance. Typically, a proton near to an electronegative element such as oxygen will experience less shielding by electrons and thus resonate at a lower frequency than one affected by an electropositive group. The overall result of shielding is that protons in a sample are spread over a range of characteristic chemical shifts. Proton NMR is referenced by setting the chemical shift for tetramethylsilane (TMS) at zero, with the result that most protons give a chemical shift in the range 0 to 10 parts per million (ppm).

A second way that NMR provides information on the local environment is through coupling. This is the effect where two (or more) nuclei exchange information on their spin state through the bonds between them. Coupling usually operates over a maximum of four chemical bonds, but more typically three or fewer. Such information is highly useful in determining which nuclei are closely connected to each other.

Lastly, nuclei can exchange information through space, sensing the presence of other nuclei within a small distance (less than 5 angstroms [\AA]) in what is known as the nuclear Overhauser effect (NOE).

The combination of these three effects allows us to gain a huge amount of information on the environments of protons in a molecule, and this has led to NMR (along with mass spectrometry) being the leading analytical technique for chemical synthesis.

NMR is not limited to hydrogen, and although some atoms do not produce an NMR response, both carbon and nitrogen have isotopes with suitable nuclei (^{13}C and ^{15}N , respectively), both of which are NMR active and behave in a similar way to protons. However, there are some important distinctions, not least the low natural abundance of ^{13}C and ^{15}N nuclei, and their intrinsic NMR characteristics. For carbon, the ^{13}C isotope is around 1.1% abundant, so that typically a ^{13}C spectrum is already 100-fold weaker than the proton spectrum for the same sample. For nitrogen the situation is worse, with an abundance of ^{15}N only 0.4%. On top of this, the relative NMR response compared with a proton is poor for both ^{13}C and ^{15}N , and thus natural abundance experiments on these nuclei are very difficult. In the context of proteins this means that isotopically enriched samples are desired and are usually prepared by expressing the protein in bacteria, with the isotopic labels introduced by feeding the bacteria specific nutrients. In most cases this will



Figure 1. A 500 MHz NMR spectrometer (Material and Chemical Characterization—MC², University of Bath)

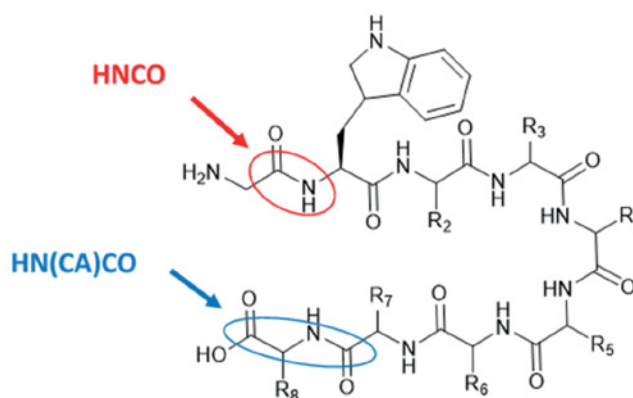
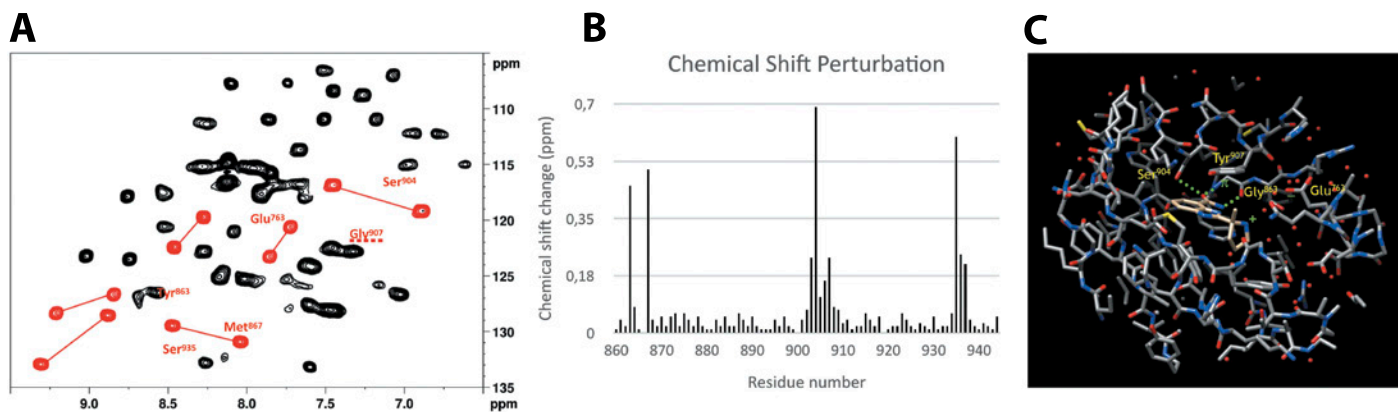


Figure 2. The HNCO experiment relates proton, carbon and nitrogen nuclei in the backbone of a protein, whereas the HN(CA)CO experiment links the carbonyl carbon nuclei of two adjacent residues in the protein by their relationship to the nitrogen and hydrogen nuclei in the amide bond. Combining these, and other, experiments allows for the assignment of all the proton, carbon and nitrogen nuclei within the protein.

Figure 3. Changes in the ^1H - ^{15}N correlation spectrum on ligand binding (chemical shift perturbation) can be used to identify residues most affected by the ligand, and thus are relevant to the binding mode. In **(A)** a ^1H - ^{15}N correlation spectrum is depicted, with the red spots highlighted as showing the greatest change in chemical shift. In **(B)** a 'skyscraper' plot shows the changes in chemical shift for all the residues (in this case from 860 to 950). In **(C)** the overall structure with affected residues highlighted.



for such a 2D spectrum is greatly reduced. For proteins this 2D spectrum provides a 'fingerprint' and is often used to ensure that the protein has folded correctly, and to check that it has not degraded.

In order to more fully assign the spectrum, and thus work out which proton signals arise from which amino acid in the chain, a 'triple resonance' approach is required. This involves using ^1H , ^{13}C and ^{15}N in the same experiment, passing magnetization from one nucleus to another, and ultimately detecting the ^1H signals. As a result of coupling (typically through one and two chemical bonds), the proton spectrum is modulated and thus the chemical shifts of the nearby carbon and nitrogen nuclei can be assigned. In practice this requires several different triple resonance experiments. Typically, one experiment, known as HNC0, will link the proton,

be via the minimal medium method, which contains all the salts and trace elements required by the bacteria, but with no carbon or nitrogen sources other than those enriched with ^{13}C and/or ^{15}N .

Proteins are complex molecules, starting with the primary, covalently bonded amino acid sequence, and extending into the secondary structure, including α -helices and β -pleated sheets that are held in shape by hydrogen bonds between residues and electrostatic interactions. These secondary structures are further ordered into the tertiary structure, where the full shape of the protein emerges, driven in part by the pressure of water molecules forcing hydrophobic areas to the internal regions of the folded protein. Lastly, single proteins routinely assemble with one or more other proteins in the quaternary structure.

For even the smallest protein, the 1D ^1H spectrum is far too complex to assign in isolation; by convention, the smallest protein has at least 30 residues and thus overlap between signals renders accurate assignment impossible. The easiest route to resolving overlap is to use a second dimension, for instance combining ^1H and ^{15}N spectra, and showing a peak in the spectrum where a proton is related by one chemical bond to a nitrogen. The overlap

nitrogen and carbon within an amino acid, whereas an HN(CA)CO correlates the proton and nitrogen with the carbonyl carbon from both its own base and the previous residue. Combining such experiments (and there are many variations) allows the full assignment of the spectrum.

The next step requires establishing the spatial arrangement of the protons in the sample, and this makes use of the NOE, as NOE contacts arise where protons are within 5 Å. All NOE contacts in a sample can be measured with a 2D experiment called NOESY—any cross peaks in the spectrum reveal that the two protons responsible are within 5 Å. The data from NOE contacts can be converted into a set of distance restraints, and these are combined with further information on bond angles derived from coupling constants, dipolar contacts and hydrogen bonds. Combining all the restraints allows computational answers to be generated that lead to an ensemble of possible structures—if enough data has been used the structures will converge.

Aside from the potential for the investigation of proteins where X-ray data is not possible, NMR also allows the protein to be studied in a more physiological state, i.e. in solution. In addition, NMR is able to

provide unique information on dynamic processes and intermolecular interactions that cannot be accessed in the more static environment of an X-ray suitable crystal. However, the NMR approach is limited by the sheer complexity of larger proteins, and also by the need for relatively large amounts of pure protein.

The use of structure-based drug design (SBDD) is widespread in the search for target-based novel small molecules. While X-ray crystallography has typically been the preferred technique, not least for its intrinsic higher resolution of ligand-protein structures, there are many cases where NMR is indispensable. Most obvious is where a given protein simply refuses to provide X-ray suitable crystals, and this includes a new class of intrinsically disordered proteins that are incompatible with crystal structure determination. However, other

problems can occur—for instance binding sites in crystals can be artefacts, created by crystal contacts, and examples where ligand binding engenders significant changes in the binding site are likely not to be accessible via soaking crystals with potential ligands.

NMR can play a significant role in SBDD at several different levels. At the top level of information, a high-resolution structure can be determined using NOE contacts, including those to the small molecule bound in the active site. Where ligand binding induces structural changes partial or complete determination of the target structure may be required.

At a lower level, protein-observed NMR measurements have been used, and this makes use of the small changes in chemical shift of nuclei near to the binding site. This method, also known as chemical shift perturbation, is predicated on the idea that when a ligand binds to a protein, the nuclei nearest to the binding site will be most affected, changing their chemical shifts. Nuclei away from this binding site will be all but unaffected. Titration experiments, where the 2D ^1H - ^{15}N correlation spectra of the protein is recorded in the absence and presence of drug allows the binding sites to be probed without recourse to full structure determination, assuming that the resonance assignments are known. As such, many ligands can be introduced to the protein to screen for those that have a significant binding interaction, and also for the location of that binding.

The lowest level involves actively screening non-labelled protein for potential binding interactions. These ligand-based screening methods make use of the NMR spectrum of the ligand, and how it is changed by binding to the much larger protein. Several NMR characteristics are crucially dependent on molecular size—notably relaxation rates (the time taken for nuclear spins to return to equilibrium after being disturbed) and the NOE effect, which is positive for molecules of over 3,000 Da, and negative for those less than 1,000 Da. When a small molecule binds to a protein its NMR characteristics change to reflect that of the protein. This makes it possible to screen for new drug candidates using unlabelled protein, merely by observing the NMR of the drug. In addition, it is possible to use more than one drug at a time, and only those that bind to the protein will register as being affected.

Over the last 15 years a radical new approach has been developed—fragment-based drug design (FBDD), and NMR has played a significant role. Here, rather than attempting to find whole molecule drug candidates, much smaller fragments are used. This has several clear advantages. Firstly, using a reduced collection of very small compounds is much more efficient than attempting to screen using full drug-like molecules, where the number of possible compounds increases exponentially with increased molecular size.

Secondly, larger molecules have a greater chance of an unfavourable interaction—smaller fragments are more able to fit into binding sites and will generate many more hits. Successful fragments can subsequently be ‘merged’, ‘linked’ and grown to give candidate molecules.

A brief introduction such as this can only scratch the surface of the field. Continuous efforts are being made to improve the sensitivity of the technique and find new ways to ask questions of the samples under study. A good example of this is advances in magnet technology. In the early 1990s a 600 MHz magnet was considered state of the art. In 2015 JEOL installed a 1.02 GHz spectrometer, and the race is on to establish working systems at 1.1 and 1.2 GHz. It is certain that NMR spectroscopy will continue to play a critical role in structural biology. ■

Further reading

- Keeler, J. (2010) *Understanding NMR Spectroscopy*, Wiley, Chichester, UK
- Cavanagh, J., Fairbrother, W.J., Palmer III, A.G., Rance, M. and Skelton, N.J. (2007) *Protein NMR Spectroscopy – Principles and Practice*. Elsevier Academic Press, USA
- Rule, G.S. and Hithcens, T.K. (2006) *Fundamentals of Protein NMR Spectroscopy*, Springer, The Netherlands
- Carneiro, M.G., AB, E., Theisgen, S. and Siegal, G. (2017) NMR in structure-based drug design, *Essays Biochem* **61**, 485–493
- Loeffler, P.A., Cuneo, M.J., Mueller, G.A., DeRose, E.F., Gabel, S.A. and London, R.E. (2011) Structural studies of the PARP-1 BRCT domain. *BMC Struct. Biol.* **11**, 37–47
- Erlanson D.A. (2011) Introduction to Fragment-Based Drug Discovery in *Fragment-Based Drug Discovery and X-Ray Crystallography* (Davies T. and Hyvönen, M., eds). Topics in Current Chemistry, vol. 317. Springer, Berlin, Heidelberg, Germany
- Jordan, J.B., Poppe, L., Xia, X. et al. (2012) Fragment based drug discovery: practical implementation based on ^{19}F NMR spectroscopy. *BMC Struct. Biol.* **55**, 678–687



Timothy J. Woodman, Department of Pharmacy and Pharmacology, University of Bath. Tim obtained his PhD from the University of Warwick (1998). Following a Royal Society Fellowship at the University of Auckland (New Zealand) and postdoctoral appointments at the universities of Leeds and East Anglia, he was appointed as NMR Spectroscopist at Bath in 2005. He was promoted to Senior NMR Spectroscopist in 2013. In addition to NMR, Tim's chief research interests include drug delivery via the skin, prostate cancer and the science of chillies. Email: t.woodman@bath.ac.uk