PROTEINS: Structure, Function, and Genetics 19:273–276 (1994)

FUTURE DIRECTIONS

Where Is NMR Taking Us?*

Angela M. Gronenborn and G. Marius Clore
Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

INTRODUCTION

The last few years have seen a quantum jump both in the size and accuracy of protein structures that can be solved by NMR.[1] These newest advances in the field are following right on the heels of the earlier development of NMR techniques for three-dimensional structure determination[2,3] resulting in a substantial number of solution structures of small proteins over of the last decade. Currently it is possible to determine the structures of proteins up to the 15–25 kDa range at a resolution comparable to 2 Å resolution crystal structures.[4] This perspective aims at highlighting where current and future developments are taking place, both in the technological arena and in regard to applications to the biological sciences.

LARGER PROTEINS

The power of NMR for structure determination over other spectroscopic techniques results from the fact that every proton gives rise to an individual resonance in the spectrum which can be resolved and identified by higher dimensional (i.e., 2D, 3D, and 4D) techniques. The measurement of nuclear Overhauser effects (NOEs) between any of these protons yields distance restraints, which in conjunction with experimentally determined torsion angle restraints permit the calculation of three-dimensional structures using one or more of a number of algorithms.[3,5,6] The final high resolution solution structures are usually derived in an iterative fashion, starting with a low resolution structure based on a subset of the experimental restraints and subsequent incorporation of all data, resolving possible ambiguities step by step.

The conventional 2D NMR-based sequential assignment methodology[2,3] which has been applied so successfully to proteins of less than 100 residues breaks down for larger proteins.[1] The spectral complexity is such that 2D experiments no longer suffice, and it is essential to increase the spectral resolution by increasing the dimensionality of the spectra. In some cases it is still possible to apply basically the same strategy as previously employed in 2D experiments, but using heteronuclear (^{15}N or ^{13}C) edited 3D experiments to increase the spectral resolution. In many cases, however, numerous ambiguities still remain and it is necessary to adopt a sequential assignment strategy based solely on well defined heteronuclear scalar couplings.[7] These experiments are now commonly used for spectral assignment of proteins above the 15 kDa range with a 28-kDa protein currently (June 1994) setting the record.[8,9] While initially time consuming, it is now possible to decrease the time requirement for 3D experiments considerably, using pulse field gradients to eliminate undesired coherence transfer pathways.[9]

While the panoply of 3D heteronuclear experiments is sufficient for the purposes of spectral assignment, yet further increases in resolution are required for the reliable identification of NOE through-space interactions. The resolving power of 4D heteronuclear-edited NOE spectroscopy is the essential feature for studying larger proteins, with 4D ^{13}C–^{15}N and ^{13}C–^{13}C edited NOESY spectra becoming more and more indispensable for the determination of high quality structures.

BETTER STRUCTURES

The quality of the final NMR protein structure increases as the number of restraints increase.[11,12] This progression in coordinate precision can be clearly appreciated from a comparison of early NMR structures, calculated with ca. 7 restraints per residue and an average atomic root-mean-square deviation (rmsd) of individual structures about the mean of 1.5 Å for backbone atoms and state of the art structures calculated with ca. 20 restraints per residue exhibiting rmsd values of 0.4 Å for the backbone atoms. Efforts to increase the precision and accuracy of NMR structures have taken different routes: more accurate interproton distance restraints derived from relaxation matrix analysis of

Address reprint requests to Angela M. Gronenborn or G. Marius Clore, Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892.

NOE intensities have been advocated,[13,14] direct refinement against NOE intensities has been proposed,[15] but, most importantly, maximizing the number of structurally useful interproton distance and torsion angle restraints, even if they are specified as loose ranges, will result in improved structures.[12] We therefore have the means to determine high resolution, high quality solution structures by *NMR*, albeit not in an easy and quick fashion. It is necessary to extract all the small NOEs in order to increase the number of restraints, in particular those that are at or just above the noise level and may be easily obscured by artifacts. This is a painstaking task and some people may opt for the more readily obtainable lower resolution structures. A good rule of thumb for assessing the quality of an NMR structure is to examine the number of restraints per residue; if the structure was calculated on the basis of 16 or more restraints per residue, it is most likely that a high quality structure will have been obtained. There is, however, no guarantee and additional measures, such as an NMR $R_{1/6}$ value in intensity based refinement[16] or complete cross validation (free R factor),[17] should be examined.

JOINT X-RAY AND NMR REFINEMENT

It is clear from the preceding discussion that NMR stands side by side with X-ray crystallography for determining high resolution structures of small to medium sized proteins. Recently a new method has been developed in which the NMR and X-ray data are combined and used simultaneously in the structure refinement.[18] It was shown that a model can readily be generated from a joint NMR/X-ray refinement which is compatible with the data from both techniques.

The implications of the joint NMR/X-ray refinement method to structural biology are of considerable significance. In particular, the full potential and future use of the method will be for structure determinations of multidomain proteins, for which only low resolution X-ray data for the entire protein are available but for which detailed structural information may be obtained by NMR on the individual domains. Using the joint X-ray/NMR refinement approach in such cases will open the way to the study of proteins which may otherwise never be structurally accessible by either of the two methods alone.

PARTIALLY FOLDED OR DENATURED PROTEINS

Interest in the structural characterization of denatured or partially folded states of proteins has increased substantially over the last few years. In part this increased attention stems from the necessity to describe the folding pathways of proteins as completely as possible, i.e., to determine the number, the order, and properties of kinetic folding intermediates. The development of pulsed hydrogen exchange NMR[19,20] has led to the detection and characterization of folding intermediates and partially folded states for a variety of proteins,[21] and important contributions with regard to time constants, structure, and stability, as well as to the question of sequential versus parallel pathways have been made by NMR. In addition, attempts at describing the structural features of unfolded proteins are becoming successful. Pockets of residual structure have been identified in denatured proteins[22,23] (as defined by other spectroscopic techniques, such as circular dichroism or fluorescence spectroscopy), suggesting that the "denatured" state may not be a truly random state. More progress in this area will most certainly be forthcoming, given the fact that resonance assignments for "denatured" proteins[23] can be achieved using labeled proteins and the above mentioned heteronuclear, multidimensional techniques which have proven so successful in studying larger proteins.

PROTEIN DYNAMICS

It is a long recognized fact that NMR spectroscopy is ideally suited to study dynamic processes. Given the notion that protein function is intrinsically linked to protein mobility, any data related to the rates and amplitudes of internal motions in proteins enhances our understanding of proteins, both at the fundamental physical level with regard to the analysis of atomic motions, as well as in assessing whether certain motions interfere or enhance particular biological functions. The measurement of heteronuclear spin relaxation rates has yielded information about backbone dynamics for a number of proteins, exploiting the availability of ^{15}N-labeled proteins.[25,26] In order to extract similar information about side chain motions, ^{13}C relaxation measurements are necessary. In principle, such measurements can be carried out for uniformly ^{13}C-labeled proteins, if care is taken to avoid complications from carbon–carbon couplings or cross-correlation. A final challenge will be to interpret all the extracted relaxation parameters in terms of a physical model for the underlying motions, eventually leading to a match between the experimentally observed motions and those predicted by molecular dynamics simulations. It seems reasonable to expect that progress in that area will be forthcoming in the next few years.

TECHNOLOGICAL ADVANCES

At present most NMR laboratories engaged in structural and dynamic studies of proteins utilize spectrometers with operating frequencies of 500 and 600 MHz for protons. Higher field spectrometers are just becoming available, with 750 MHz instruments offered and under installation by commercial manufacturers. Quests for Gigaherz spectrometers are

ongoing in the United States at the National High Magnetic Field Laboratory and at Battelle, Pacific North West Laboratories. These higher field instruments will increase the resolution and sensitivity of NMR spectroscopy, so that submillimolar samples will become amenable to structural investigations. Likewise, the number of observable NOEs (and therefore structural restraints) for a given protein sample will no doubt increase with increased sensitivity, allowing even better structures to be determined by NMR. In addition, pulsed field gradient technology shows considerable impact in solvent signal and artifact supression[10] as well as coherence order selection,[27] rendering heteronuclear 3D experiments less time consuming and more sensitive.

Advances in hardware are being paralleled by progress in data acquisition, as well as processing and interpretation software. Software development is fuelled by the demands created by the plethora of different multidimensional experiments, enormously increased data sizes, and more complicated processing schemes. Alternatives to fast Fourier transforms such as Bayesian analysis, maximum entropy reconstruction, and linear prediction have been proposed. Likewise schemes for automated resonance and NOE assignments are becoming available. This latter, however, is presently still an area where user expertise is the most valuable commodity and some user intervention will probably always be necessary.

CONCLUDING REMARKS

The recent development of a whole range of highly sensitive multidimensional heteronuclear edited and filtered NMR experiments has revolutionized the field of protein structure determination by NMR. Proteins and protein complexes in the 20 kDa range are now amenable to detailed structure analysis in solution, and current methods can probably be extended to systems even up to 40 kDa providing that they are well behaved from an NMR perspective. Nevertheless, despite these advances, it should always be borne in mind that there are a number of key requirements that have to be satisfied to permit a successful structure determination of larger proteins and protein complexes by NMR. The protein in hand must be soluble and should not aggregate up to concentrations of about 1 mM, it must be stable at room temperature or slightly higher for considerable time periods, it should not exhibit significant conformational heterogeneity that could result in extensive line broadening, and finally it must be amenable to uniform ^{15}N and ^{13}C labeling. At the present time there are only a few examples in the literature of proteins in the 15–25 kDa range that have been solved by multidimensional heteronuclear NMR spectroscopy. One can anticipate, however, that over the next few years, by widespread use of multidimensional heteronuclear experiments coupled with semiautomated assignment procedures, many more NMR structures of proteins and protein–ligand complexes will become available.

ACKNOWLEDGMENTS

We thank our many colleagues, past and present, who have contributed to the work carried out in our laboratory. Above all, we thank Ad Bax with whom we have shared numerous stimulating discussions, fruitful experiments, and a continuous and most enjoyable collaboration in the best of scientific spirits.

REFERENCES

1. Clore, G.M., Gronenborn, A.M. Structures of larger proteins in solution: Three- and four-dimensional heteronuclear NMR spectroscopy. Science 252:1390–1399, 1991.
2. Wüthrich, K. NMR of Proteins and Nucleic Acids. New York: Wiley, 1986.
3. Clore, G.M., Gronenborn, A.M. Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. CRC Crit. Rev. Biochem. Mol. Biol. 24:479–564, 1989.
4. Clore, G.M., Gronenborn, A.M. Comparison of the solution nuclear magnetic resonance and X-ray crystal structures of human recombinant interleukin-1β. J. Mol. Biol. 221: 47–53, 1991.
5. Havel, T.F., Wüthrich, K. An evaluation of the combined use of nuclear magnetic resonance and distance geometry for the determination of protein conformation in solution. J. Mol. Biol. 182:281–294, 1985.
6. Braun, W. Distance geometry and related methods for protein structure determination from NMR data. Q. Rev. Biophys. 19:115–157, 1987.
7. Bax, A., Grzesiek, S. Methodological advances in protein NMR. Acc. Chem. Res. 26:131–138, 1993.
8. Fogh, R.H., Schipper, D., Boelens, R., Kaptein, R. ^{1}H, ^{13}C, and ^{15}N NMR backbone assignments of the 269-residue serine protease PB92 from *Bacillus alcalophilus*. J. Biomol. NMR 4:123–128, 1994.
9. Remerowski, M.L., Domke, T., Groenewegen, A., Pepermans, H.A.M., Hilbers, C.W., van de Ven, F.J.M. ^{1}H, ^{13}C and ^{15}N NMR backbone assignments and secondary structure of the 269-residue protease subtilisin 309 from *Bacillus lentus*. J. Biomol. NMR 4:257–278, 1994.
10. Bax, A., Pochapsky, S.S. Optimized recording of heteronuclear multi-dimensional NMR spectra using pulsed field gradients. J. Magn. Reson. 99:638–643, 1992.
11. Havel, T.F. An evaluation of computational strategies for use in the determination of protein structure from distance constraints obtained by nuclear magnetic resonance. Prog. Biophys. Molec. Biol. 56:43–78, 1991.
12. Clore, G.M., Robien, M.A., Gronenborn, A.M. Exploring the limits of precision and accuracy of protein structures determined by nuclear magnetic resonance spectroscopy. J. Mol. Biol. 231:82–102, 1993.
13. Borgias, B.A., James, T.L. COMATOSE, a method for restraint refinement of macromolecular structure based on two-dimensional nuclear Overhauser spectra. J. Magn. Reson. 79:493–512, 1988.
14. Boelens, R., Konig, T.M.G., Kaptein, R. Determination of biomolecular structures from proton-proton NOEs using a relaxation matrix approach. J. Mol. Struc. 173:299–311, 1988.
15. Yip, P., Case, D.A. A new method for refinement of macromolecular structures based on nuclear Overhauser effect spectra. J. Magn. Reson. 83:643–648, 1991.
16. James, T.L. Relaxation matrix analysis of two dimensional nuclear Overhauser effect spectra. Curr. Opin. Struc. Biol. 1:1042–1053, 1991.
17. Brünger, A.T., Clore, G.M., Gronenborn, A.M., Saffrich, R., Nilges, M. Assessing the quality of solution nuclear magnetic resonance structures by complete cross validation. Science 261:328–331, 1993.
18. Shaanan, B., Gronenborn, A.M., Cogen, G.H., Gilliland, G.L., Veerapandian, B., Davies, D.R., Clore, G.M. Combining experimental information from crystal and solution

studies: Joint X-ray and NMR refinement. Science 257: 961–964, 1992.
19. Roder, H. Elöve, G.A., Englander, S.W. Structural characterization of folding internediates in cytochrome c by H-exchange labelling and proton NMR. Nature (London) 335: 694–699, 1988.
20. Udgaonkar, J.B., Baldwin, R.L. NMR evidence for an erly framework intermediate on the folding pathway of ribonuclease A. Nature (London) 335:700–704, 1988.
21. Woodward, C.K. Hydrogen exchange rates and protein folding. Curr. Opin. Struc. Biol. 4:112–116, 1994.
22. Neri, D., Billeter, M., Wider, G., Wüthrich, K. Determination of residual structure in a urea denatured protein, the *434* repressor. Science 257:1559–1563, 1993.
23. Lumb, K.J., Kim, P.S. Formation of a hydrophobic cluster in denatured bovine pancreatic trypsin inhibitior. J. Mol. Biol. 236:412–420, 1994.
24. Logan, T.M., Olejniczak, E.T., Xu, R.X., Fesik, S.W. A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments. J. Biomol. NMR 3:225–418, 1993.
25. Torchia, D.A., Nicholson, L.K., Cole, H.B.R., Kay, L.E. Hetereonuclear NMR studies of the molecular dynamics of staphylococcal nuclease. In NMR of Proteins. New York: Macmillan Press, 1993:190–219.
26. Wagner, G., Hyberts, S., Peng, J.W. Study of protein dynamics by NMR. In NMR of Proteins. New York: Macmillan Press, 1993: 220–257.
27. Kay, L.E., Keifer, P., Saarinen, T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J. Am. Chem. Soc. 114:10663–10665, 1992.