PROTON MAGNETIC RESONANCE SPECTROSCOPIC STUDIES USING PARAMAGNETICS OF PRIMARY AND TERTIARY STRUCTURE OF PROTEINS

J. H. BRADBURY, LARRY R. BROWN, M. W. CROMPTON and BRIAN WARREN

Chemistry Department, Australian National University, Canberra, ACT, Australia

ABSTRACT

On binding a paramagnetic broadening probe at the end of a random coil polypeptide chain, one obtains sequential broadening of the α-CH resonances along the chain and, hence, can determine the sequence of the peptide. The method has been applied to tripeptides (0.25 mg of material being used), tetrapeptides and a hexapeptide under conditions such that gadolinium ions bind at the C-terminus or cupric ions at the N-terminus.

Dimethylation of the lysine residues of lysozyme and observation of the chemical shifts of the six dimethyl proton resonances as a function of pH allows the determination of their pK. From studies of selective broadening of the methyl resonances on addition of gadolinium ions and the known structure of the lysozyme–ion complex, it is possible to obtain the pK values of the lysine residues of lysozyme.

Approaches to the determination of the structure in solution of proteins using paramagnetic broadening and shifting probes is discussed. Problems concerning the site of ion binding and the assignment of p.m.r. resonances arise when the crystal structure is unknown.

The amount by which a proton magnetic resonance (p.m.r.) line is broadened by a paramagnetic probe with a long electronic relaxation time —e.g. gadolinium(III)—is dependent on the inverse sixth power of the distance between the nucleus giving rise to the resonance and the paramagnetic probe. Furthermore, if the paramagnetic probe is an ion which binds rapidly and reversibly at a specific site, then the width of the resonance line will progressively increase as the concentration of paramagnetic ion is increased until such time as the site is saturated, after which no further broadening will occur.

On the binding of gadolinium ions at low concentration to lysozyme, there is preferential broadening of p.m.r. resonances that arise from protons which are closest to the binding site[1,2]. The binding site has been shown by x-ray studies[3] to be located between glutamic acid 35 and aspartic acid 52. By the use of difference spectroscopy[4–6], in this case the subtraction of a paramagnetically broadened spectrum from the spectrum of native lysozyme, it has been possible to observe the resonances from the side chains of valine

109 and alanine 110[1,2]. In addition, distance information has been obtained from these studies with broadening probes and also from the use of lanthanide ions with short electronic relaxation times (europium(III), praseodymium(III)), which cause shifts of resonances[2,7]. In this paper three applications are given of the use of paramagnetic ions in peptide and protein chemistry.

1. DETERMINATION OF PRIMARY STRUCTURE (SEQUENCE) OF PEPTIDES

If a peptide is present as a random coil in aqueous solution (which can be achieved by use of a denaturant such as urea or guanidine hydrochloride), then the average distance from one end of the molecule to successive α-CH groups along the peptide chain increases as one proceeds along the chain. Thus, addition of a small amount of a suitable paramagnetic ion which is bound at one end causes most broadening of the α-CH resonance of the nearest amino acid residue, less of the second residue and very little of the third residue[8]. This effect is shown in *Figure 1*, for addition of gadolinium

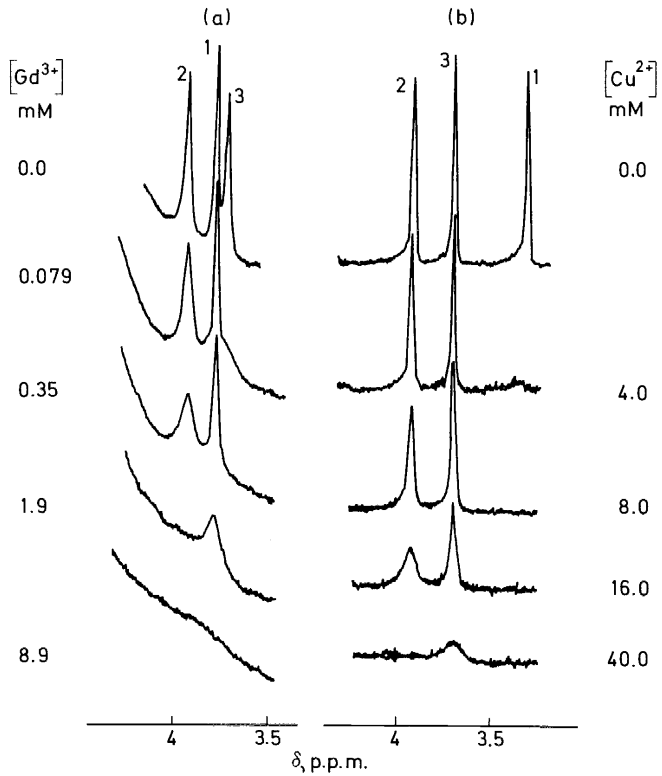

Figure 1. The p.m.r. spectra at 100 MHz of the methylene protons of triglycine numbered sequentially from the N-terminal end (assignment due to Sheinblatt[9]): (a) 2% solution of triglycine in 8 M urea at pH 4.3 on addition of Gd^{3+} at the molar concentrations shown; (b) 4% solution of triglycine in D_2O at pH 11, using Cu^{2+}

ion which binds at the C-terminal carboxylate site and broadens the resonances in the expected sequence 3. 2. 1. Also, cupric ion binds at the N-terminus and causes broadening of residues in the expected sequence 1. 2. 3. It is noted that the differences in the positions of the numbered resonances from (a) to (b) are due to shifts of the resonances which result from the change in pH of the solution[9].

Triglycine is a simple tripeptide to examine because the CH_2 resonances have double the intensity of the α-CH resonances of other amino acids and also because the latter are split into multiplets. A more difficult case is the tetrapeptide Gly–Leu–Gly–Leu, the results of which are given in *Figures 2* and *3*. The N-terminal glycyl CH_2 resonance and the C-terminal leucyl α-CH resonance are readily identified by the manner in which they shift with

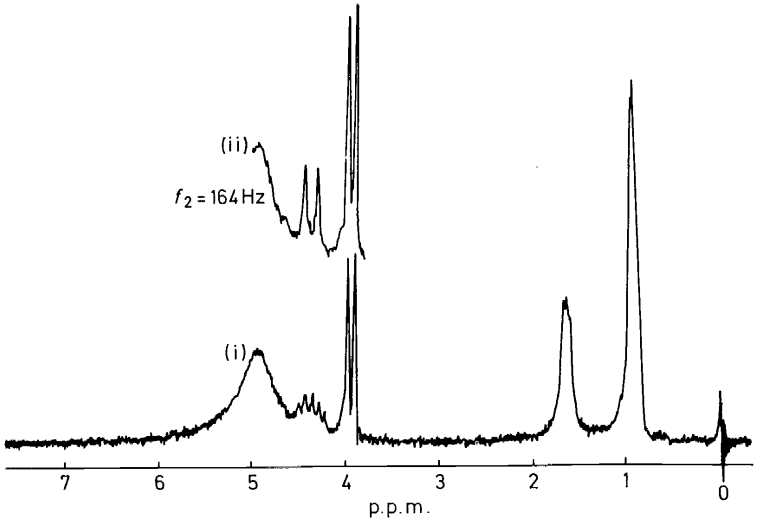

Figure 2. The p.m.r. spectra at 100 MHz of 4% solution of Gly–Leu–Gly–Leu in 8 M urea at pH 4.3: (i) showing two overlapping leucyl α-CH triplets at about 4.3 p.p.m., two CH_2 glycyl resonances at 4 p.p.m. and the β-CH_2 leucyl resonances at 1.6 p.p.m.; (ii) showing the collapse of the α-CH leucyl triplets to singlets owing to irradiation of the β-CH_2 leucyl resonances at 1.6 p.p.m.

change of pH[9]. Irradiation of the β-CH_2 resonances of the leucyl residues at 1.64 p.p.m. causes the collapse of the α-CH triplets into singlets as shown in *Figure 2*. The HDO resonance is small owing to repeated lyophilization of the peptide + urea at pH 4.3 from D_2O and use of 100.0% D_2O (Aldrich Chemical Co.). It is also conveniently broad at this pH because of an intermediate rate of exchange of protons between D_2O and deuterated urea.

The results obtained from the broadening of the resonances on addition of gadolinium ions are shown in *Figure 3*. Under double resonance conditions with all resonances occurring as singlets, the correct order of broadening—4. 3. 2. 1—is obtained. In the absence of decoupling, it is found that the resonance from leucine 2 does not decrease in height (owing to partial decoupling from the β-CH_2) over a range of concentrations for which the resonance from glycine 1 falls off more rapidly. However. it is noted that the

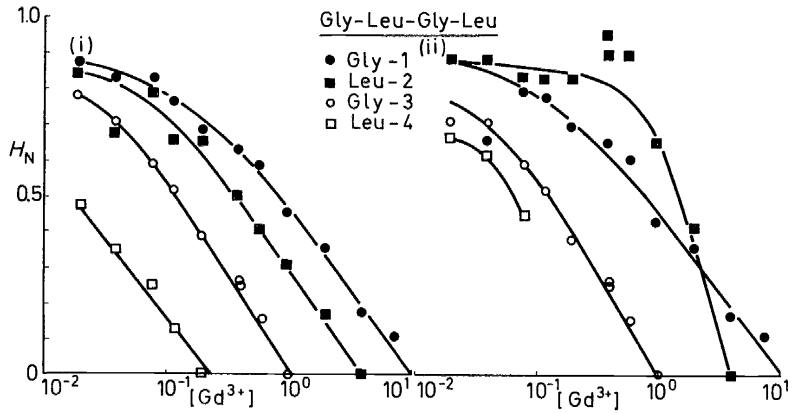

Figure 3. Graph of H_N (height of the α-proton resonances of Gly–Leu–Gly–Leu normalized to 1 in the absence of Gd^{3+}) against $[Gd^{3+}]$ in mM for (i) spin decoupled leucyl α-CH resonances (*Figure 2*(ii)) and (ii) normal undecoupled system (*Figure 2*(i))

correct order of broadening is established at higher concentrations of gadolinium ions and is followed if one considers the point of complete disappearance of the peaks. Clearly it is desirable to use a broad band double resonance technique, such that all the α-CH resonances are converted to singlet resonances, thus giving the advantages shown of sharper resonances and a more readily identifiable sequence of broadening.

Some difficulty has been encountered with large HDO resonance peaks which tend to obscure some α-CH resonances. This is shown in *Figure 4*, where the triplet of the α-CH seryl resonance at 4.6 p.p.m. is partly obscured by the broad HDO resonance in the upper spectrum. The alanyl α-CH quartet is located at about 4.2 p.p.m. and the glycyl CH_2 singlet at 4.0 p.p.m. with part of the seryl β-CH_2 resonance visible on the upfield side. As in the case of *Figure 2*, the HDO resonance is small owing to repeated lyophilization of the peptide + urea from D_2O and broad because of the intermediate rate of exchange of protons between D_2O and deuterated urea. The lower spectrum is obtained by irradiation at the frequency f_2, which causes saturation of the HDO resonance[10].

A second problem concerns the possibility of the binding of gadolinium ions on side chain carboxyl groups and of cupric ions on side chain nitrogen atoms. The former is prevented by working at pH 4.3, since experiments with Glu–Val–Phe have shown that there is no appreciable amount of binding of gadolinium ions to the side chain carboxyl group. This presumably results from the fact that the latter is a weaker acid than the α-carboxyl group; hence, the side chain carboxyl is largely protonated, whereas the α-carboxyl is unprotonated at pH 4.3. Experiments with dipeptides and tripeptides with lysine, arginine and histidine residues in the C-terminal position have shown that at pH 8 in 6 M guanidine hydrochloride there is no appreciable amount of binding to the lysyl or arginyl side chains. Unfortunately there is a considerable degree of binding on the histidyl side chain, and work is in progress on this problem.

A third question concerns the probable length of peptides which can be

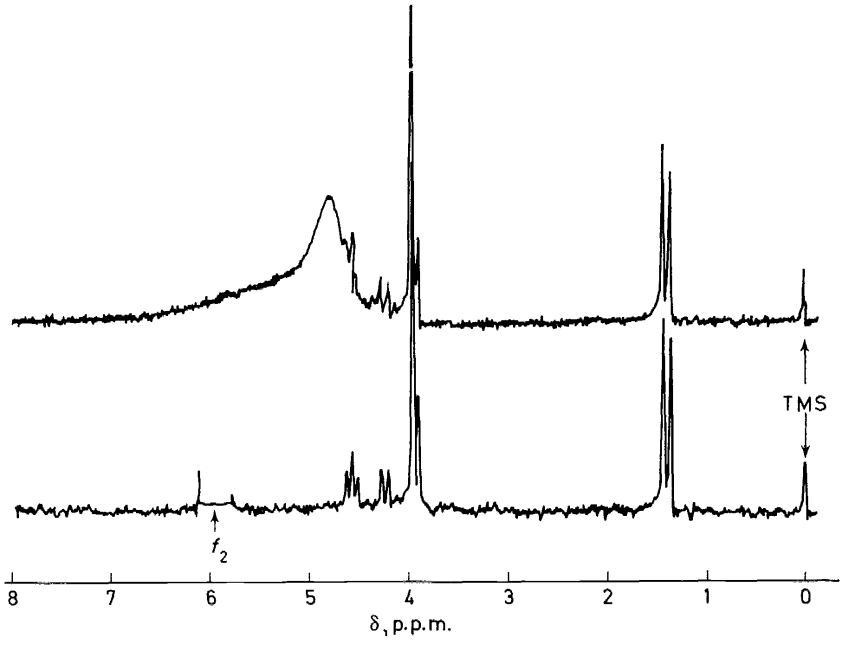

Figure 4. The p.m.r. spectra at 100 MHz of a 2% solution of Gly–Ser–Ala in 8 M urea at pH 4.3 in D_2O, with and without irradiation at f_2

sequenced by this procedure. Experiments with hexaglycine have shown that this peptide can be readily spanned either from the C-terminus with gadolinium ions or from the N-terminus with cupric ions. It should therefore be possible to sequence a peptide with possibly 12 residues, although another limitation on length is imposed by the complexity of the p.m.r. spectrum.

Finally, the amount of peptide required for sequencing has been reduced to about 0.25 mg (for a tripeptide at 100 MHz) by use of a micro-tube of 0.05 ml capacity and a small amount of spectral accumulation[8].

2. DETERMINATION OF DISSOCIATION CONSTANTS OF LYSINE RESIDUES IN LYSOZYME[11]

The ε-amino groups of lysine residues in proteins can be methylated by treatment with formaldehyde and borohydride at 0°[12]. When this procedure is applied to a number of proteins, it is found from amino acid analysis that the extent of dimethylation of lysine residues varies from 66% with bovine serum albumin to 92% for ribonuclease. In lysozyme there are six lysine residues, and the amino acid analysis shows that 4.1 residues are dimethylated and 0.4 residues are N-ε-monomethylated. Since separate proton resonances can be observed from each of the N-dimethyl lysine groups, it is clear that all six lysine residues in lysozyme are reacted.

The *N*-methylated lysozyme retains complete enzymatic activity, and observation of its p.m.r. spectrum as a function of temperature in D_2O

shows that its heat denaturation is almost identical with that of lysozyme. Again, the p.m.r. spectra of the native and methylated lysozyme have been compared closely at 100 MHz by difference spectroscopy and in the high field region at 270 MHz. The spectra of the native and modified proteins are almost the same except for the differences introduced by the *N*-dimethyl lysine residues. It is therefore concluded that native lysozyme and methyl lysozyme have very similar, if not identical, three-dimensional structures in solution.

When the methyllysozyme is titrated with base and the pH is raised from 8 to 12, there are progressive upfield shifts of the *N*-dimethyl resonances numbered 1–6 in *Figure 5*, due to deprotonation of the *N*-dimethyl groups. The peak labelled 7 in *Figure 5* broadens and disappears above pH meter reading 6.6. It appears to be related to peak 6 and there is some evidence that peaks 6 and 7 both refer to the same lysine residue (probably lysine 1), which may have two different environments in the molecule[11]. It is possible

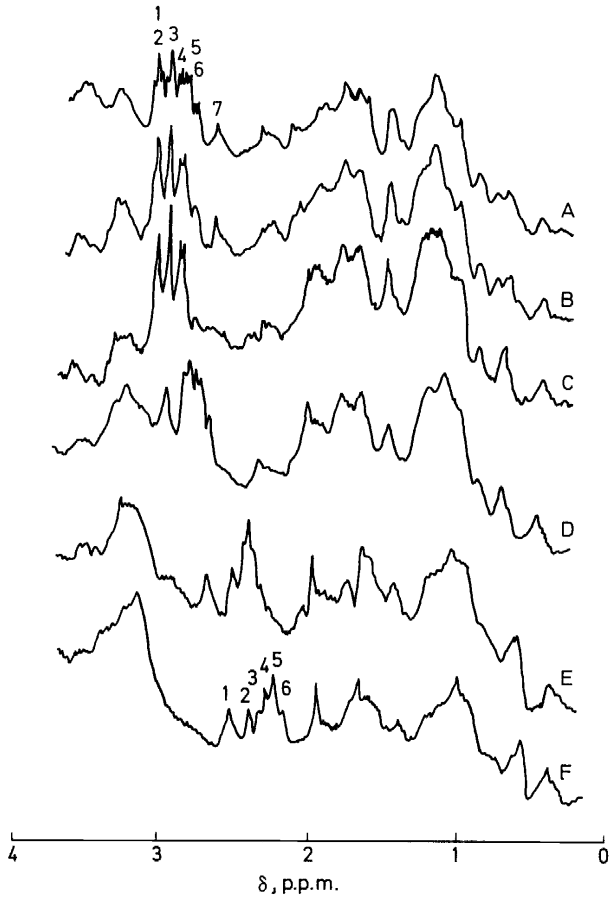

Figure 5. The p.m.r. spectra at 270 MHz of the *N*-methyl resonances of methyllysozyme (2% solution in D_2O) at pH meter reading: A, 3.08; B, 5.08; C, 6.66; D, 9.14; E, 10.17; and F, 10.60

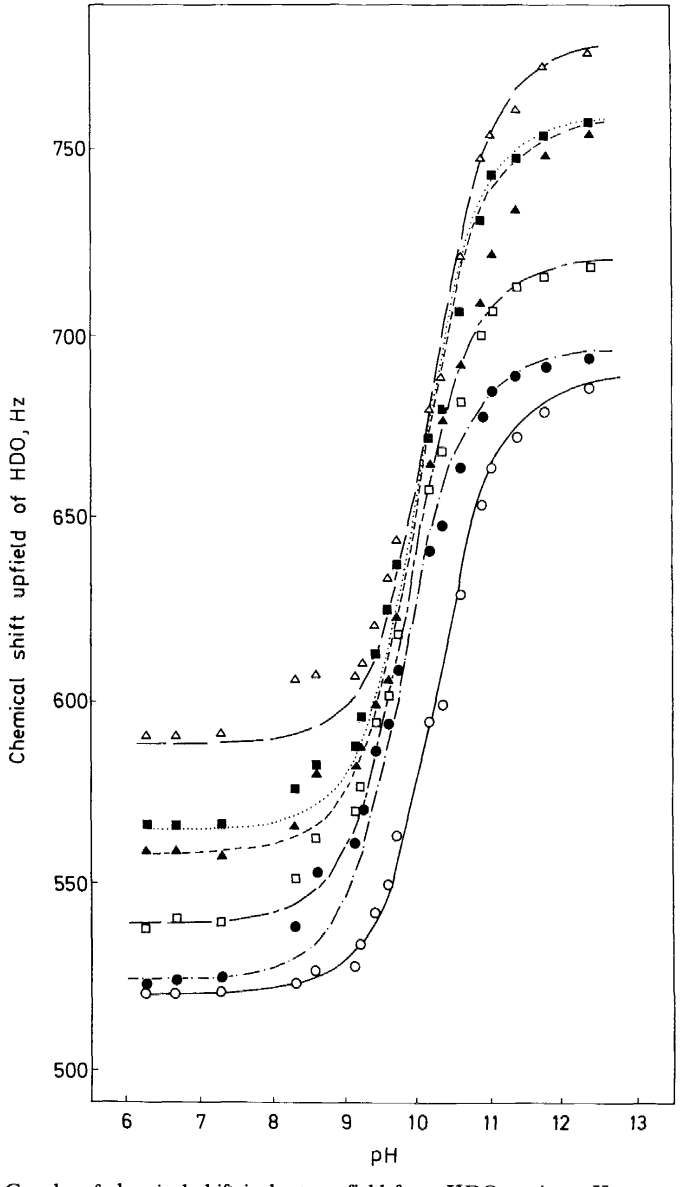

Figure 6. Graphs of chemical shift in hertz upfield from HDO against pH meter reading for *N*-dimethyl resonances of methyllysozyme at 270 MHz and 18°

to follow the peak position of each resonance as a function of pH and, hence, construct the family of titration curves shown in *Figure 6*. These have been analysed as simple, non-interacting ionizations by use of the Henderson–Hasselbach equation and pK values have been determined by a least squares analysis. The range of pK values of 9.6–10.2 refers to *N*-dimethylamino side chains at 18°. A small temperature correction is required to 25° and an

increase of 0.6 (based on studies with model compounds[11]) to obtain reasonable values for amino side chains (see *Table 1*).

The assignment of resonances 1–6 to specific lysine residues in methyllysozyme is achieved by the selective broadening of these resonances on

Table 1. p*K* values of the lysine residues of lysozyme at 25°C

Residue number in sequence	p*K*	Number of resonance (see *Figure 5*)	Residue number in sequence	p*K*	Number of resonance (see *Figure 5*)
1	10.6	6	96	10.7	1
13	10.3	4	97	10.1	2
33	10.4	5	116	10.2	3

addition of gadolinium ions. The site of gadolinium binding in lysozyme has been established by x-ray diffraction studies on the crystal[3] and two reasonable assumptions are made: (1) that the solution structure of lysozyme is the same as that in the crystal[2] and (2) that the solution structures of lysozyme and methyllysozyme are the same (see above). It is thus possible to calculate the distances from the gadolinium binding site to the various *N*-methyl protons and then compare these with the order in which broadening of resonances 1–7 occurs, as shown in *Figure 7*. It is noted that the chemical shifts of resonances 1 and 2 are identical at this pH and, hence, their broad-

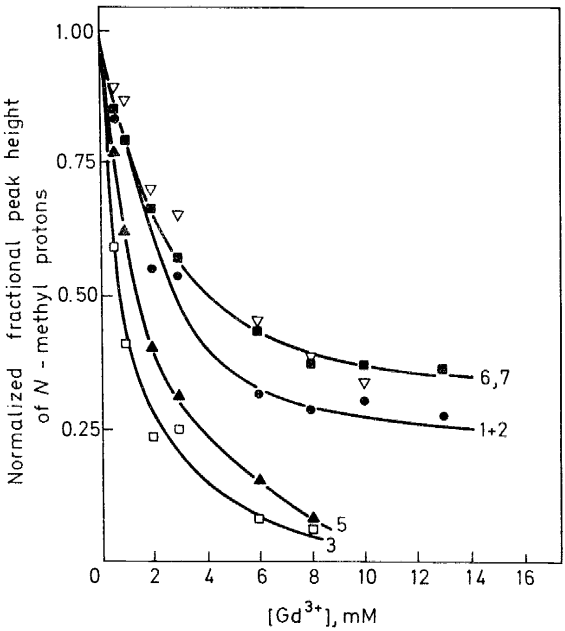

Figure 7. Plot of fractional peak height (normalized against peak 4, which shows very little decrease in height until high gadolinium concentrations and has been used as an internal standard) against $[Gd^{3+}]$ at pH 4.5 in D_2O

ening behaviour could not be separated. However, it is possible to make a tentative assignment between them on the basis of their known environment (from x-ray studies) and their appreciable differences in pK[11]. The final assignments are given in *Table 1*, together with the corrected pK values (which are thermodynamic apparent pK values) for the various lysine residues in lysozyme.

3. APPROACHES TO DETERMINATION OF TERTIARY STRUCTURE OF PROTEINS

Previous work in this field using p.m.r. spectroscopy and broadening and shifting lanthanide probes[2] has been concerned with its application to a protein of known crystal structure (lysozyme) in which the site of lanthanide binding is also known. This work is of fundamental importance because of the need to check the validity of the p.m.r. procedures (applied in solution) using a known structure (determined in the crystal). This check is of course dependent on an assumption with regard to the similarity or otherwise of the structure in the crystal and in dilute solution. The evidence obtained from the lysozyme study[2] would appear to indicate that the p.m.r. procedures are valid and also that the structure of the protein in solution is very closely the same as in the crystal.

In application of broadening and shifting probes to p.m.r. spectroscopic studies of the structure of a protein of known sequence and unknown tertiary structure, there are two difficult problems to be resolved. Firstly, the site(s) of binding of the lanthanide ions along the polypeptide chain needs to be known. One method of achieving this which is currently under investigation is the incorporation of a synthetic binding site for lanthanides at a known site in the protein. Secondly, it is necessary to assign as many p.m.r. resonances as possible to *specific* protons in the molecule. Detailed assignments are usually difficult to achieve because there is normally more than one residue of a particular type per molecule. Thus in bovine α-lactalbumin there are three histidine residues and, hence, three C-2 protons and three C-4 protons which give rise to resonances. It has been possible to assign these resonances by chemical means[13], but it is worth noting that the few assignments made in the absence of a knowledge of the tertiary structure have involved chemically reactive groups such as histidine[14, 15] and tryptophan[16]. It appears that improved methods are required for the solution of these two problems (and perhaps other problems too), before the n.m.r. technique will be able to make a substantial contribution in its own right towards determination of protein structure in solution.

ACKNOWLEDGEMENTS

We wish to thank Drs R. E. Richards and I. D. Campbell for providing access to the facility of the 270 MHz p.m.r. spectrometer at Oxford University and Professor D. C. Phillips for provision of laboratory facilities during a period of sabbatical leave by J.H.B. in 1972. The Australian Research Grants Committee is thanked for the provision of financial support for this work.

REFERENCES

[1] I. D. Campbell, C. M. Dobson, R. J. P. Williams and A. V. Xavier, *J. Magn. Resonance*, **11**, 172 (1973).
[2] I. D. Campbell, C. M. Dobson, R. J. P. Williams and A. V. Xavier. *Ann. NY Acad. Sci.* **222**, 163 (1973).
[3] C. C. F. Blake and M. A. Rabstein, personal communication.
[4] B. Bak, E. J. Pedersen and F. Sundby, *J. Biol. Chem.* **242**, 2637 (1967).
[5] N. L. R. King and J. H. Bradbury, *Nature (Lond.)* **229**, 404 (1971).
[6] J. H. Bradbury and N. L. R. King, *Austral. J. Chem.* **25**, 209 (1972).
[7] C. D. Barry, A. C. T. North, J. A. Glasel, R. J. P. Williams and A. V. Xavier, *Nature (Lond.)* **232**, 236 (1971).
[8] J. H. Bradbury, M. W. Crompton and Brian Warren. *Analyt. Biochem.* **62**. 310 (1974).
[9] M. Sheinblatt, *J. Amer. Chem. Soc.* **88**, 2845 (1966).
[10] J. Feeney and G. C. K. Roberts, *Chem. Commun.* 205 (1971).
[11] J. H. Bradbury and Larry R. Brown, *Europ. J. Biochem.* **40**, 565 (1973).
[12] G. E. Means and R. E. Feeney, *Biochemistry*, **7**, 2192 (1968).
[13] R. S. Norton, *Ph.D. Thesis*, Australian National University (1974).
[14] D. H. Meadows, O. Jardetzky, R. M. Epand, H. Rüterjans and H. A. Scheraga, *Proc. Nat. Acad. Sci., Wash.* **60**, 766 (1968).
[15] J. H. Bradbury and B. E. Chapman, *Biochem. Biophys. Res. Commun.* **49**, 891 (1972).
[16] J. D. Glickson, W. D. Phillips and J. A. Rupley, *J. Amer. Chem. Soc.* **93**, 4031 (1971).