N.M.R. Studies on Myelin Basic Protein. I
$^{13}$C Spectra in Aqueous Solutions

Bogdan E. Chapman and Walter J. Moore
School of Chemistry, University of Sydney, N.S.W. 2006.

Abstract

Carbon-13 n.m.r. spectra have been obtained for bovine myelin basic protein at pD 4.4 in D$_2$O and in 6 M guanidine deuterochloride solutions. Chemical-shift differences between resonances from some amino acid residues are interpreted in terms of structured regions in the polypeptide chain of the native protein, whereas the denatured protein displays the spectrum expected for an essentially random coil. Measurements of $T_1$ and n.O.e. provide quantitative data on the dynamics of the backbone and side-chain carbons, and give support to the conclusion that the native protein does not have a random-coil structure.

Introduction

The conformation in solution of myelin basic protein has been the subject of several recent investigations using the technique of nuclear magnetic resonance spectroscopy.$^{1-3}$ These studies, however, have not yielded quantitative information on the structure of the protein or on the molecular motions within the protein. Such information is desirable since the mobility of various domains in the protein structure, as well as its average conformation, can influence the functional properties of the protein, in particular its activity in cell-mediated immune processes.$^{4,5}$

The purpose of the present work has been twofold: (i) to obtain experimental $^{13}$C n.m.r. spectra that would help to elucidate the structure of the protein in aqueous solution; (ii) to investigate the internal motions of the polypeptide chain of this unusual protein in both its native and denatured forms.

Myelin basic protein (MBP) has several unusual properties. With a molecular mass of 18500 daltons, it has no conventional secondary structure, such as an $\alpha$-helix or a $\beta$-pleated sheet, in aqueous solution.$^6$ As the pH of the solution is raised from 4.0 to 7.0 and above, the protein forms dimers and possibly higher oligomers, although the exact proportions of the various oligomers have not yet been elucidated.$^7,8$

The primary structures of the human, bovine, and rat proteins have been determined. On the basis of the unusual triprolyl (PPP) sequence at residues 99-101, Brostoff and Eylar suggested that a 'hairpin bend' occurred in the polypeptide chain with the PPP triplet acting as a hinge. Theoretical calculations of the conformation of the PPP sequence indicated a minimum of potential energy of 40 kJ/mol at torsional angles $\phi = 165^\circ$, $\phi' = 165^\circ$. This theoretical conformation is similar to the suggested hairpin bend.

High-resolution n.m.r. spectroscopy is most useful in the study of molecular motions in proteins since different regions of the protein molecule can be observed simultaneously. Internal motions in polymers were detected early by this method. More recent work has focused on detailed studies of the intramolecular motions of specific residues, mainly aromatic. It has been shown that the rotational motion of these groups varies from being restricted, with non-equivalence of chemical shift, to being sufficiently fast that chemical-shift non-equivalence due to tertiary structure is averaged out.

As the data from a single relaxation parameter can often be satisfied by several models, it is necessary to measure at least two different n.m.r. relaxation parameters to obtain quantitative values of the effective correlation times of individual nuclei in amino acid residues. Two or more relaxation parameters satisfied by one correlation time will drastically restrict the number of models compatible with the experimental data.

Fourier-transform $^{13}$C n.m.r. spectroscopy is especially useful for the study of molecular motions. With conventional instruments it is possible to measure two relaxation parameters, the longitudinal relaxation time ($T_1$) and the nuclear Overhauser enhancement (n.O.e.), values of which are determined by the intramolecular motions of the protein. The dependence of these two parameters on correlation time is sufficiently different to permit the elimination of many models for mobility of the polypeptide chain and its side chains. In principle, a third parameter, the transverse relaxation time ($T_2$), can also be measured, although there are experimental difficulties with large molecules. Estimates of $T_2$ can often be obtained, however, from the widths of spectral lines.

The theoretical interpretation of the $^{13}$C data is simplified by the fact that most of the carbon atoms have directly bonded hydrogens, so that their dominant relaxation mechanism should be dipolar. For isotropic motion, accurate correlation times can therefore be obtained, since the distance between dipoles is given by the C-H bond length. Theoretical treatments are available also for some anisotropic motions.

11 Dunkley, P. R., and Carnegie, P. R., Bioch. J., 1974, 141, 243.
such as the rotational diffusion of an ellipsoid and the rotation of a side chain superimposed on the rotation of the molecule as a whole. In a protein structure, we may expect carbon nuclei of a given type (e.g., α carbons of polypeptide chains) to display a distribution of correlation times for intramolecular motions, and hence the distribution function is an additional variable that must be considered in the interpretation of dynamic data.

**Experimental**

Myelin basic protein was isolated from bovine brain white matter according to the method of Eylar et al. The protein was purified by passage through a Sephadex-G75 (2.5 by 75 cm) column eluted by HCl at pH 2. The purity of the protein was checked by sodium dodecyl sulfate–polyacrylamide gel electrophoresis. When 20 µg of purified protein was applied to 15% gels in 0.2% sodium dodecyl sulfate, only a single strong band was observed on staining with Coomassie Blue. Although the MBP is certainly purified sufficiently for n.m.r. studies, chemical evidence indicates that certain impurities, notably neutral proteinase activity, are still present in small concentrations.

Measurements of the 13C spectra, longitudinal relaxation times and nuclear Overhauser enhancements were performed on a Bruker HX-90 Fourier-transform pulsed spectrometer operating at 22.625 MHz.

Samples were run in 15 mm o.d. sample tubes. The protein concentration was usually 10−2 M in D2O at pH 4.4 (meter reading ±0.4). The probe temperature was 30°C. Each spectrum was obtained by averaging 30000 transients at 5000 Hz spectral width, 1 Hz digital broadening and 8192 time domain points. A 14-µs (90°) pulse was used to excite the 13C nuclei. Protons were decoupled from the carbon nuclei by means of a Bruker broadband decoupler. Dioxan was used as an internal standard for measurement of chemical shifts, which are reported in ppm from external Me2Si (67.85 ppm upfield from dioxan).

The spectrum in 6 M guanidine deuter chloride was obtained in 10 mm o.d. tubes, a Bruker HXQ50 quadrature detection receiver being used. Quadrature detection gives an improvement in signal to noise by a factor of 1.4. Sample and experimental conditions were as above.

Longitudinal relaxation times were measured by means of the inversion–recovery technique, with the pulse sequence (180°−τ−90°−DE), with τ the variable delay time between the two pulses and DE the delay interval between each acquisition which is set to five times the longest T1 being measured (0.8 s in this instance). Nine values of τ were used ranging from 0.02 to 4.0 s.

Nuclear Overhauser enhancement was measured by the gated decoupling technique. The proton decoupler is turned on at the same time that a 90° pulse is used to excite the 13C nuclei. At the end of acquisition of the free induction decay from the 13C nuclei, the decoupler is turned off and a delay of eight times the longest T1 of interest is set before the next sequence of pulses. The experiment results in a completely decoupled spectrum without n.O.e. The ratios of the line intensities as given by the peak areas to those in a continuously decoupled spectrum give the value of n.O.e.

Spectra at higher fields were obtained at 67.87 MHz on a Bruker HX-270 n.m.r. spectrometer, 10 mm o.d. sample tubes being used, at the National N.M.R. Centre, Canberra. Sample conditions were as before and instrumental conditions are given in the figure legends.

**Results**

**Spectral Assignments**

The proton-decoupled natural abundance 13C n.m.r. spectrum of bovine myelin basic protein at 22.625 MHz in D2O at a concentration of 10−2 M at pH 4.4 and

---

Fig. 1. $^{13}$C natural abundance spectra of myelin basic protein in D$_2$O at pH 4.4 and 30°C. Spectra at 22.625 MHz.
(a) $10^{-2}$ M in D$_2$O.
(b) $10^{-2}$ M in 6 M guanidine deuterocloride.
(c) Simulated spectrum in 6 M guanidine deuterocloride based on parameters in Table 1.
N.M.R. Studies on Myelin Basic Protein. I

303 K is shown in Fig. 1a. The spectrum under the same conditions but in solution in 6 M guanidine deuterochloride is shown in Fig. 1b. The backbone and side-chain carbonyl resonances appear between 170 and 185 ppm downfield from tetramethylsilane. Arginine and tyrosine ζ resonances appear at 158 and 156 ppm. Resonances between 110 and 140 ppm are from aromatic side-chain carbons. Peaks from α carbons appear between 50 and 65 ppm. The upfield region from 10 to 45 ppm

Table 1. Parameters for 13C n.m.r. spectrum of myelin basic protein in 6 M guanidine deuterochloride in D2O at pD 4.4

<table>
<thead>
<tr>
<th>Carbon</th>
<th>Chemical shift</th>
<th>Carbon</th>
<th>Chemical shift</th>
<th>Carbon</th>
<th>Chemical shift</th>
<th>Carbon</th>
<th>Chemical shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>51.25^A</td>
<td>α</td>
<td>43.70^D</td>
<td>x</td>
<td>54.29^A</td>
<td>x</td>
<td>56.59^A</td>
</tr>
<tr>
<td>β</td>
<td>17.83^b</td>
<td>β</td>
<td>31.18^b</td>
<td>β</td>
<td>28.97^b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C'</td>
<td>177.22^c</td>
<td>γ</td>
<td>30.44^e</td>
<td>γ</td>
<td>110.37^c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td></td>
<td></td>
<td>15.52^b</td>
<td>δ</td>
<td>126.02^e</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>55.03^a</td>
<td></td>
<td>57.14^a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>29.06^b</td>
<td></td>
<td>37.07^d</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ</td>
<td>25.84^e</td>
<td></td>
<td>129.43^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>δ</td>
<td>42.04^f</td>
<td></td>
<td>129.43^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ζ</td>
<td>157.97^f</td>
<td></td>
<td>129.43^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C'</td>
<td>176.48^c</td>
<td></td>
<td>129.43^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp^-</td>
<td>51.71^A</td>
<td></td>
<td>25.65^e</td>
<td>C'</td>
<td>177.40^c</td>
<td></td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>37.62^d</td>
<td></td>
<td>49.23^b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ (CO)</td>
<td>177.03^c</td>
<td></td>
<td>132.19^g</td>
<td>Pro</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C'</td>
<td>174.46^e</td>
<td></td>
<td>117.09^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp^-</td>
<td>53.74^A</td>
<td></td>
<td>25.47^e</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>40.66^b</td>
<td></td>
<td>49.23^b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ (CO)</td>
<td>179.80^c</td>
<td></td>
<td>132.19^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C'</td>
<td>175.56^c</td>
<td></td>
<td>117.09^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>55.39^A</td>
<td></td>
<td>55.03^a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>β</td>
<td>27.49^b</td>
<td></td>
<td>62.67^b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ</td>
<td>32.47^g</td>
<td></td>
<td>129.43^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>53.92^A</td>
<td></td>
<td>132.19^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C'</td>
<td>177.03^c</td>
<td></td>
<td>132.19^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>33.27^h</td>
<td></td>
<td>132.19^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>55.39^A</td>
<td></td>
<td>132.19^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C'</td>
<td>176.21^c</td>
<td></td>
<td>132.19^g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^A Width 8.0 Hz, relative n.O.e. 3.0. ^B Width 5.0 Hz, relative n.O.e. 3.0. ^C Width 2.0 Hz, relative n.O.e. 2.1. ^D Width 12.0 Hz, relative n.O.e. 3.0. ^E Width 5.0 Hz, relative n.O.e. 3.0. ^F Width 2.0 Hz, relative n.O.e. 2.3. ^G Width 4.0 Hz, relative n.O.e. 3.0.
consists of signals from methylene and methyl groups. These general assignments are based on tables of standard $^{13}$C chemical shifts for amino acid residues in proteins.\(^{28,29}\)

\begin{figure}
\centering
\includegraphics[width=\textwidth]{spectrum.png}
\caption{$^{13}$C natural abundance spectrum of myelin basic protein (10^{-2} M) in D$_2$O at pD 4.4 and 30°C recorded at 67.9 MHz. Sweep width 15 kHz, repetition time 2.16 s, 65536 time domain addresses, 93670 scans. (a) Region of carbonyl resonances; (b) region of aliphatic resonances. The region of aromatic resonances is shown in Fig. 3.}
\end{figure}

Refinement of the assignments was made by a computer simulation of the spectrum. The program used was similar to that of Lilley et al.\(^{30}\) with minor modifications.


Fig. 3. Region of aromatic resonances in the natural abundance $^{13}\text{C}$ n.m.r. spectrum of myelin basic protein recorded at 67.9 MHz under the conditions described in Fig. 2.
The carbons in each type of amino acid residue were assigned characteristic frequencies, linewidths and n.O.e. values. The line shapes were fitted to Lorentzian functions and the spectrum was plotted as the sum of these lines, each weighted by the frequency of occurrence of the amino acid residue in the primary structure of the protein. By visual comparison of the simulated and experimental spectra, modifications of frequencies and widths were introduced until a satisfactory fit was obtained. The resulting simulated spectrum for the protein denatured in 6 M guanidine deuterchloride is shown in Fig. 1c. The frequencies and linewidths employed in the final simulation are summarized in Table 1.

The experimental spectrum in 6 M guanidine deuterchloride is simulated accurately by the calculated spectrum in the aromatic regions and in the methyl group regions at 10-20 ppm. The experimental spectrum in D$_2$O, however, displays definite discrepancies with the simulation, especially in the resonances centred about 16.5 and 19.5 ppm, which have been assigned to isoleucine and valine methyls, respectively. The $\alpha$ carbon and methylene regions (20-70 ppm) could not be satisfactorily simulated for the spectra of either native or denatured protein. This result is hardly surprising, however, since this region comprises overlapping peaks from different amino acid residues and it is virtually impossible to decide which carbon resonances are incorrectly represented in the simulation. No attempt was made to simulate the carbonyl region at 170-185 ppm, since the chemical shifts vary with the particular one of the 108 pairs of amino acid residues in the dipeptide units, and even to some extent with more distant interactions.

To obtain a better resolution of the spectral peaks, data were obtained at 67.9 MHz at the National N.M.R. Laboratory (Canberra). These spectra for the protein in D$_2$O are shown in two sections of Fig. 2 and in Fig. 3. Fig. 2a shows the region of the carbonyl resonances, Fig. 2b shows the region of $\alpha$ carbons and side-chain methyl and methylene groups, and Fig. 3 shows the region of aromatic side chains.

The carbonyl region in Fig. 2a displays 38 distinct peaks. It is obvious that these cannot be simulated by 23 different individual carbonyl resonances ascribed to the amino acid residues in MBP. The available assignments of peptide-bond carbonyls are fragmentary. It is probable that the three downfield peaks at 178.6 ppm are the $\delta$ carbons of glutamic acid ($N = 2$) and glutamine ($N = 8$). Much additional experimental data from small peptides will be required before detailed analysis of the carbonyl spectrum of a protein such as MBP will be feasible.

The region of aromatic side chains in Fig. 3 displays a number of interesting features. Most noteworthy is the chemical-shift splitting of the phenylalanine (F) $\gamma$ peaks at 137.4 ppm. At least two distinct lines can be distinguished spanning a width of 30 Hz (Fig. 4b). These non-protonated F$\gamma$ carbons would be the most likely to display resolution of chemical-shift heterogeneities. The protonated F$\zeta$ peak shows traces of heterogeneity. On the other hand, the tyrosine (Y) $\gamma$ peak at 129.1 ppm is split into an unequal doublet. The assignments of F$\zeta$ and Y$\gamma$ were made definitely on the basis of their measured $T_1$ values, 150 and 700 ms respectively, since the protonated carbons must relax more rapidly. None of the other F or Y peaks shows chemical-shift splitting despite the high resolution of the spectrum. There are eight phenylalanine and four tyrosine residues in the protein.

The 17 arginine (R) residues yield only a single sharp R$\zeta$ peak in this region.

The histidine (H) peaks are characteristically broad as would be expected from the lability of the ionizable protons even at pD 4.4.
The individual tryptophan (W) peaks from the sole tryptophan residue in the protein are all detected, except for Wε3, which is missing under the Hδ envelope. Assignment of these W peaks was facilitated by comparison with the spectrum of a synthetic sample of the peptide encephalitogenic in guinea pig, FSWGAEGQK, comprising residues 115–122 in the bovine protein. The chemical-shift differences between peptide and protein are shown in Table 2.

<table>
<thead>
<tr>
<th>Carbon</th>
<th>Peptide</th>
<th>Protein</th>
<th>Δδ</th>
<th>Carbon</th>
<th>Peptide</th>
<th>Protein</th>
<th>Δδ</th>
</tr>
</thead>
<tbody>
<tr>
<td>γ</td>
<td>109.90</td>
<td>110.26</td>
<td>+0.36</td>
<td>η2</td>
<td>123.27</td>
<td>123.16</td>
<td>−0.11</td>
</tr>
<tr>
<td>ε2</td>
<td>113.23</td>
<td>113.16</td>
<td>−0.07</td>
<td>δ1</td>
<td>125.80</td>
<td>125.46</td>
<td>−0.34</td>
</tr>
<tr>
<td>ε3</td>
<td>119.49</td>
<td>119.59</td>
<td>+0.10</td>
<td>δ2</td>
<td>128.18</td>
<td>128.96</td>
<td>+0.78</td>
</tr>
<tr>
<td>ε3</td>
<td>120.68</td>
<td>120.36</td>
<td>−0.32</td>
<td>ε2</td>
<td>137.51</td>
<td>137.18</td>
<td>−0.33</td>
</tr>
</tbody>
</table>

Fig. 4. Comparison of spectra at 67.9 MHz. (a) Protein denatured in 6 M guanidine deuterochloride (conditions as in Fig. 2, except 186800 scans); (b) native protein in D2O.

High-field spectra of MBP in 6 M guanidine deuterochloride could not be obtained with the same signal-to-noise ratio and resolution as those of the protein in D2O, owing to difficulties in tuning the probe and in controlling the temperature of the sample. Expanded spectra of the phenylalanine γ peak and of the upfield methyl region for the denatured protein are shown in Fig. 4a. These spectra were the best obtained in six attempts. In the spectrum of denatured protein, the closely spaced components of the Fδ peak have collapsed to a single broad peak. Also (not shown) the splitting of the Yδ peak disappears on denaturation.

The tentative assignments in the α carbon region from 64 to 42 ppm may be found in Table 1. With the exception of the glycine (G) α peaks at 44 ppm, they are not at this stage supported by definitive evidence.

The region from 45 to 22 ppm consists of methylene peaks with the exception of leucine (L) methyls at 23·7 and 22·2 ppm. The assignments in Table I are not definite.

The region of the methyl peaks from 20 to 14 ppm contains some interesting chemical-shift splittings. In particular the threonine (T) δ resonance at 20·4 ppm is split into two peaks. There are seven threonine residues in the protein. The valine (V) methyl resonance at 19·5 ppm appears as a triplet (instead of the expected doublet due to non-equivalence of methyl groups). There are only three valine residues in the protein. Upon denaturation in 6 M guanidine deuterohloride the splitting of the T and V methyl resonances collapses as both then display homogeneity in their chemical shifts.

The region from 15 to 17 ppm in Fig. 4b shows three peaks. The peak at 15·6 ppm is from methionine γ-methyl carbons and the two peaks at 16·1 and 16·8 ppm are from isoleucine methyl carbons and show chemical-shift heterogeneity. In 6 M guanidine deuterohloride/D₂O most of this heterogeneity is removed and the isoleucine resonances occur at 15·8 ppm, the position expected in a random coil. There is still a small peak at 16·5 ppm which would indicate that MBP retains some residual structure even in this strong denaturant.

Measurements of $T_1$ and N.O.E.

Typical experimental spectra that were used to calculate $T_1$ values are shown in Fig. 5. The areas under the various peaks were estimated in two ways, by the product of peak height and half-width and by the weight of the cut-out graph paper for the expanded peaks. The signal intensities $I(\tau)$, proportional to the peak areas, were plotted as log($I_{\infty} - I$) against delay time $\tau$, but the $T_1$ values were computed from the slope of the best least-squares fit of the linear regression line

$$\ln(I_{\infty} - I) = \ln(2I_{\infty}) - \tau/T_1$$

Here $I_{\infty}$ is the equilibrium intensity. The errors in the individual $I$ values were assumed to be inversely proportional to $I$ and corresponding weights were applied to the data terms in the least-squares program. The $T_1$ values were also calculated by the method of Gerhards and Dietrich. The two methods gave results that were concordant within 4%. The values of $T_1$ so calculated are included in Table 3 together with their standard deviations.

It should be noted that these $T_1$ values do not correspond to individual carbons but must be regarded as an average over the $T_1$ values of the carbons included in the peak in question. For example, the $T_1$ values for the $\alpha$ carbons were based upon four different peaks, in addition to an integrated intensity of the envelope of the entire $\alpha$ carbon region.

In a few instances, the $I$ against $\tau$ data were fitted to the sum of two exponential decay functions:

$$I_{\infty} - I = I_A \exp(-\tau/T_{1A}) + I_B \exp(-\tau/T_{1B})$$

Since the standard deviations of the monophasic $T_1$ values were at worst 10%, these biphasic decay curves were close to the limit of statistical significance. Nevertheless

---

Table 3. Values of $T_1$ and n.O.e. for myelin basic protein (pD 4.4) at 22.625 MHz

$[\text{mbp}] 10^{-2} \text{ m}$, except where stated otherwise

<table>
<thead>
<tr>
<th>Peak</th>
<th>Solvent</th>
<th>$T_1$ (ms)</th>
<th>S.d. (ms)</th>
<th>n.O.e.</th>
<th>Peak</th>
<th>Solvent</th>
<th>$T_1$ (ms)</th>
<th>S.d. (ms)</th>
<th>n.O.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$-CA</td>
<td>$D_2O$</td>
<td>80</td>
<td>±8</td>
<td>2.0</td>
<td>Tyr $\varepsilon$</td>
<td>$D_2O_b$</td>
<td>162</td>
<td>±8</td>
<td>2.2</td>
</tr>
<tr>
<td>$\alpha$-CA</td>
<td>6 m guanidine deuterochloride</td>
<td>69</td>
<td>±11</td>
<td>2.3</td>
<td>Ile CH$_3$</td>
<td>$D_2O_b$</td>
<td>310</td>
<td>±10</td>
<td>2.0</td>
</tr>
<tr>
<td>$\alpha$-CA</td>
<td>$D_2O_b$</td>
<td>—</td>
<td>—</td>
<td>2.0</td>
<td>Ala CH$_3$</td>
<td>6 m guanidine deuterochloride</td>
<td>—</td>
<td>—</td>
<td>2.5</td>
</tr>
<tr>
<td>Ala CH$_3$</td>
<td>$D_2O_b$</td>
<td>300</td>
<td>±8</td>
<td>2.2</td>
<td>Phe $\varepsilon$</td>
<td>6 m guanidine deuterochloride</td>
<td>137</td>
<td>±8</td>
<td>2.3</td>
</tr>
<tr>
<td>Phe $\varepsilon$</td>
<td>$D_2O_b$</td>
<td>133</td>
<td>±7</td>
<td>2.2</td>
<td>Phe $\gamma$</td>
<td>6 m guanidine deuterochloride</td>
<td>—</td>
<td>—</td>
<td>1.7</td>
</tr>
<tr>
<td>Phe $\gamma$</td>
<td>$D_2O_b$</td>
<td>1130</td>
<td>±120</td>
<td>1.7</td>
<td>Tyr $\varepsilon$</td>
<td>6 m guanidine deuterochloride</td>
<td>140</td>
<td>±3</td>
<td>2.5</td>
</tr>
<tr>
<td>Ile CH$_3$</td>
<td>6 m guanidine deuterochloride</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td>6 m guanidine deuterochloride</td>
<td>140</td>
<td>±3</td>
<td>2.5</td>
</tr>
</tbody>
</table>

$^A$ Average of four peaks at 56.9, 54.9, 51.2 and 43.9 ppm. $^B$ $[\text{mbp}] 2.78 \times 10^{-3}$ m.
Fig. 6. Measurement of the nuclear Overhauser enhancement of myelin basic protein by the gated decoupling method. (a) Spectrum with n.O.e.; (b) spectrum without n.O.e.
it is likely that they reflect a genuine dispersion of the $T_1$ values. For example, for
the alanyl $\mathrm{CH}_3$

$$T_{1A} = 90 \text{ ms} \quad T_{1B} = 400 \text{ ms}$$

Spectra that were used to calculate n.O.e. values by the gated decoupling method
are shown in Fig. 6. The spectra with and without n.O.e. must be scaled to the same
noise value. The n.O.e. results are included in Table 3.

The effect of concentration of the protein on n.O.e. was not marked; this is
indicated by comparing the values for $10^{-2}$ and $2.78 \times 10^{-3}$ M solutions as given
in Table 3.

Discussion

Chemical Shifts and Linewidths

The general pattern of the $^{13}\text{C}$ spectra of MBP in $\mathrm{D}_2\mathrm{O}$ at pH 4.4 displays mainly
quite sharp peaks with a smaller number of broader peaks. This type of spectrum
would be consistent with the occurrence in the protein of some regions that have
tertiary or quaternary structure, with resultant restriction in mobility, and other
regions that are capable of less restricted internal motions as in a random-coil
polypeptide chain. Similar spectra have been observed for solutions of histone H4
in water.\(^{35}\) In the case of the histone, which is a basic protein of molecular weight
11300, there is a region of secondary structure of $\alpha$-helical type. In the case of MBP,
however, the possible structural regions do not fit any standard type of secondary
structure. Indeed the principal problem is to distinguish intramolecular folding of
the polypeptide chain from intermolecular aggregation (quaternary structure).

At pH 7.8, the peaks become markedly broadened,\(^{3}\) and independent evidence\(^{8}\)
indicates that the protein forms dimers and higher oligomers at this pH and the
concentrations used for the $^{13}\text{C}$ measurements.

When the protein is dissolved in 6 M guanidine deuterochloride, the broad peaks
become equally sharp at both pH 4.4 and 7.8. Thus this powerful denaturing agent
converts the protein into a monomeric form. Denaturation may also destroy most
of any residual non-random structure in the monomeric polypeptide chain although
the broad $F\gamma$ peak would indicate that some non-covalent interactions are still
present. The denaturing action of guanidine deuterochloride has been observed in
the cases of several native proteins with well defined globular structures, e.g.,
ribonuclease\(^ {33}\) and lysozyme.\(^ {34}\)

The $^{13}\text{C}$ n.m.r. spectra of globular proteins such as cytochrome-c and myoglobin,\(^ {17}\)
in some of their resonance peaks, display a chemical-shift heterogeneity that is
ascribed to their folded structures. In the $^{13}\text{C}$ n.m.r. spectrum of MBP the lines from
amino acid residues of identified types in most cases do not display any chemical-
shift splitting. This result is consistent with large regions of essentially random-coil
structure in the protein in aqueous solution. Chemical-shift heterogeneity is most
often observed in resonances from non-protonated aromatic carbons. These carbons
give rise to narrow lines, which occur in regions of the spectrum with minimal overlap of
resonances.


The chemical-shift splitting of certain aromatic and aliphatic carbons indicates that structured regions exist in the polypeptide chain. The definite model of the chain conformation proposed by Brostoff and Eylar\textsuperscript{12} was based on a hairpin bend at the triprolyl sequence 99–101. This bend extended at least through residues 89–114 and brought the unique methylarginine at 107 opposite the pair of phenylalanines at 89 and 90. Calculations by Scheraga \textit{et al.}\textsuperscript{13} indicated a rather deep minimum in potential energy for a folded conformation of the PPP sequence similar to that proposed by Brostoff and Eylar. A calculated potential-energy map obtained by Khanarian\textsuperscript{35} from an adaptation of the Scheraga program (QCP 286) is shown in Fig. 7.

Fig. 7. Calculated conformational energy of the triprolyl sequence. Numbers shown are energies above the global minimum at (165°,165°). The program for the calculation was written in kcal (1 kcal = 4.184 kJ) and this unit was retained for convenience.

The stabilization of the bent conformation does not depend upon hydrogen bonding and hence may be relatively resistant to disruption by guanidinium chloride. There is thus good theoretical support for a bend at the PPP sequence, but the nature of the conformation following a bend remains an open question.

The \textsuperscript{13}C n.m.r. evidence is consistent with a structured region in the neighbourhood of a PPP bend, based on the following evidence.

The sequence 85–118 contains three of the eight phenylalanine residues and none of the four tyrosine residues. The chemical-shift heterogeneity of the F\textgamma and F\textzeta carbons is the most marked among all unprotonated aromatic carbons.

One must consider the hypothesis that the splitting in the resonances might possibly be due to the primary sequence of MBP, in which four of the phenylalanine residues exist in adjacent FF pairs. That this is probably not so is indicated by work on carp muscle calcium-binding parvalbumin, which has a phenylalanine pair in the sequence but shows no splitting of the F\textgamma peak in the denatured form. Under non-denatured conditions, on the other hand, splitting was observed in the 'EF fragment peptide' which does not contain the adjacent phenylalanines of parvalbumin.\textsuperscript{36}

\textsuperscript{35} Khanarian, G., personal communication.

\textsuperscript{36} Jardetzky, O., personal communication.
There are only three valine residues in the protein, at 87, 88 and 95, all in neighbourhood of the PPP sequence. The heterogeneity of the valine resonances is clearly evident. There are three isoleucine (I) residues, at 36, 93 and 153. One Ie resonance is markedly shifted downfield, whereas the other Ie peak (the area of which corresponds to two carbons) is at a position only slightly downfield from the random-coil value. The n.m.r. data therefore support (but do not yet prove) the theoretical indications of a folded conformation in the neighbourhood of the PPP sequence.

The $^{13}$C data, which indicate the existence of structured regions in MBP, are supplemented by extensive high-resolution proton spectra, obtained from the protein in dilute solutions at low pH, in which the extent of dimerization is negligible. Analysis of these $^1$H data also indicates that at 15% concentration and pD 4.4, as used for the $^{13}$C spectra, about 5% of the protein would be in dimeric form.

Reorientation Correlation Times from $T_1$ and N.O.E. Data

The two parameters $T_1$ and n.O.E. can be interpreted in terms of overall and internal motions of the molecule. For a rigid sphere

$$1/T_1 = (1/10)(\kappa^2/\rho^6)f(\tau_R)$$

(1)

where

$$f(\tau_R) = \tau_i/[1 + (\omega_H - \omega_C)^2\tau_R^{-2} + 3\tau_i/(1 + \omega_C^{-2}\tau_R^{-2}) + 6\tau_i/[1 + (\omega_H + \omega_C)^2\tau_R^{-2}]]$$

$\kappa = h\gamma_H\gamma_C$, $\gamma_H$ and $\gamma_C$ being the gyromagnetic ratios of the proton and carbon; $\omega_H$ and $\omega_C$ are resonance frequencies of protons and carbons, $r$ is the carbon–proton distance, and $\tau_R$ is the rotational correlation time.

This equation is most applicable to protonated carbons but can also be used for non-protonated aromatic carbons if all interactions, except for those with protons two bonds away, are neglected. The above equation as written is applicable to a carbon nucleus which is relaxed by a single proton. If more than one proton is involved in relaxation of a carbon nucleus the effects must be summed

$$1/T_1 = \sum_{n=1}^{N} 1/T_{1n}$$

where $N$ is the number of protons involved.

Dynamics of Backbone Carbons of the Polypeptide

The measured $T_1$ for the $\alpha$ carbons has an average value for the several peaks of 80±5 ms. From the theroretical equation (1), based on the model of isotropic rotary diffusion of a rigid body, the corresponding values of the rotational correlation time $\tau_R$ could be either 26 or 0.7 ns.

The n.O.E. value of 2.0, on the isotropic model, would indicate $\tau_R = 2$ ns. The linewidth of 12 Hz, obtained through computer simulation of the spectrum, would suggest a $\tau_R$ of about 5 ns.

The effective $\tau_R$ values for MBP may be compared with the $\tau_R$ for the $\alpha$ carbons of poly-L-proline ($M = 11500$) of 0.61 ns, and of poly(prolylglycine) ($M = 13200$) of 0.36 ns. These values were obtained from the measured $T_1$ and equation (1)

in the extreme narrowing limit since the measured n.O.e. were maximal. In the case of ribonuclease (M = 13500) the $T_1$ value of the $\alpha$ carbons for the native protein was 42 ms, and for the protein denatured in acid (pH 1.46) it was 120 ms. The corresponding values of $\tau_R$ were estimated by combining the $T_1$ data with computer simulation of the linewidths that provided an effective $1/T_2^*$. These $\tau_R$ were 30 ns for native and 0.40 ns for denatured protein. In the case of myoglobin (M = 16900) $T_1$ measurements at two different field strengths allowed an unequivocal determination of $\tau_R$ of 22 ± 5 ns for the $\alpha$ carbons.

For muscle calcium-binding parvalbumin (M = 12000) measurements of $T_1$ and n.O.e. were combined to yield an effective $\tau_R$. For the $\alpha$ carbons, $T_1 = 55$ ns and n.O.e. = 1.2 to yield $\tau_R = 12$ ns. The structure of the native protein is constant with a rather rigid conformation of the $\alpha$ carbons, which then tumble with the overall rotational motion of the compact molecule. When MCBP was denatured by heat and guanidinium chloride, the n.O.e. of the side-chain carbons increased but no value was cited for the $\alpha$ carbons.

### Table 4. Experimental and calculated relaxation parameters for the $\alpha$ carbons of myelin basic protein (10^{-2} M, pD 4.4) at 30°C and 22.63 MHz

<table>
<thead>
<tr>
<th>Experimental parameter</th>
<th>Isotropic model</th>
<th>Ellipsoid model</th>
<th>log$\chi^2$ model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_1$ = 80 ms</td>
<td>0.7</td>
<td>4.5</td>
<td>0.8</td>
</tr>
<tr>
<td>N.O.e. = 2.0</td>
<td>2.0</td>
<td>16.0</td>
<td>1.0</td>
</tr>
<tr>
<td>$W_{1/2} = 12$ Hz</td>
<td>5.0</td>
<td>25.0</td>
<td>0.7</td>
</tr>
</tbody>
</table>

**MBP in D_2O**

| $T_1$ = 69 ms          | 0.8            | 5.5            | 0.8             |
| N.O.e. = 2.3           | 1.5            | 6.5            | 0.9             |
| $W_{1/2} = 8$ Hz       | 2.0            | 12.0           | 1.5             |

**MBP in D_2O/guanidine deuterchloride**

The backbone correlation times derived from measurements of $T_1$, peak width at half height ($W_{1/2}$) and n.O.e. for MBP in D_2O and 6 M guanidine deuterchloride are given in Table 4. The single isotropic correlation time model shows that, in D_2O, the calculated $\tau_R$ values corresponding to the n.m.r. parameters $T_1$, $W_{1/2}$ and n.O.e. vary by nearly one order of magnitude.

Evidence from viscosity and electron microscopy indicated that MBP is a rod-shaped ellipsoid with an axial ratio as high as 10. Low-angle X-ray diffraction, however, indicated a statistical-coil structure rather than a rigid ellipsoid.

An attempt was made to fit the relaxation data on MBP to the model of a rigid ellipsoid with $a = 15$ nm and $b = c = 1.5$ nm, the values estimated from viscosity data for a rod-shaped ellipsoid. These values yield a relation between $\tau_A$ for reorientation perpendicular to the $a$ axis and two additional correlation times, $\tau_B$ and $\tau_C$. The angle of the CH vector to the long axis of the ellipsoid was assumed to

---

be 109.5°. The results obtained were not much better than those from the isotropic model.

The next model is to consider the effect of a distribution of correlation times on the n.m.r. relaxation parameters as in the theory of Schaefer, who showed how the n.m.r. relaxation parameters of viscous polymer solutions can be accounted for by assuming a distribution of correlation times. Following Schaefer, we use a log $\chi^2$ distribution:

$$Fp(s)ds = [(ps)^{p-1}C^{-p}p/\Gamma(p)]ds$$

$$s = \log_b[1+(b-1)\tau/\bar{\tau}]$$

$Fp(s)ds$ is the probability of finding a correlation time $\tau$ corresponding to $s = s(\tau)$. A logarithmic time scale (of log to the base $b$) is used to allow broad distributions of correlation times. $\bar{\tau}$ is the mean correlation time of the distribution. The width of the distribution is characterized by $p$ for a fixed $b$. As $p$ increases the distribution becomes narrower approaching a $\delta$-function for very large $p$. According to Schaefer, $p$ may be regarded as a measure of the number of degrees of freedom. There is a logarithmic decrease in the width of the distribution with increasing $p$. A distribution with $p = 100$ ($b = 1000$) gives calculated relaxation parameters nearly identical with that for a single isotropic correlation time, while $p = 2$ gives a distribution of correlation times covering many orders of magnitude.

The n.m.r. relaxation parameters were calculated as a function of $\bar{\tau}$ and $p$ for $b = 1000$. Table 4 shows that a distribution width defined by $p = 15$ (a moderate distribution of correlation times) satisfies the n.m.r. relaxation parameters of MBP in D$_2$O. The variation in $\bar{\tau}$ from $T_1$, $W_{h/2}$ and n.O.e. is less than two. The value of $\bar{\tau}$, approx. 1 ns, indicates that MBP is a very flexible molecule compared to native globular proteins with $\tau_R$ values in the range 12–30 ns.

When MBP is placed in 6 M guanidine deuterochloride the distribution narrows dramatically and the relaxation parameters can be fitted well by the single isotropic correlation time model. A slight improvement can be obtained by using a log $\chi^2$ distribution with $p = 60$. In this denaturing solvent the correlation times defining the motion of the backbone polypeptide chain are approximately the same.

**Dynamics of Side-Chain Carbons**

A theoretical treatment for the rotation of C–H groups about an axis attached to a main chain which tumbles isotropically has been developed by Woessner and applied to $^{13}$C–$^1$H dipolar relaxation by Doddrell et al. The relaxation time ($T_1$) of a carbon undergoing internal rotation is determined by $\tau_R$, the overall isotropic reorientation time, $\tau_{m}$, the correlation time for rotation of the side chain, and $\theta$, the angle of the C–H vector to the axis of internal rotation. As most carbons in side chains can have many different motions, the analysis is restricted to those groups, such as the $\beta$ methylene carbons, the $\beta$ methyl of alanine and the aromatic side chains, which can be defined by a single internal rotation superimposed upon an overall isotropic rotation.

The $T_1$ and n.O.e. values for the CH$_3$ of alanyl residues are included in Table 3. There are 13 alanyl residues in MBP, including the acetylated N-terminal alanine. They all appear in an unresolved single peak at 17.83 ppm.
The $\beta$ methyl of alanine would be expected to have an angle of rotation (\(\theta\)) of 109.5°, the tetrahedral angle, for rotation about the Cx-C$\beta$ bond. The values of \(NT_1 = 900\) ms and n.O.e. = 2.2 cannot be accounted for by the isotropic model with a single correlation time. For isotropic motion \(NT_1\) gives a value for \(\tau_R\) of \(5.2 \times 10^{-11}\) or \(3.2 \times 10^{-7}\) s and the n.O.e. gives a value of \(1.5 \times 10^{-9}\) s. For a single internal motion with the mean correlation time of 1 ns obtained from Schaefer analysis for MBP in D$_2$O as the \(\tau_R\) value, \(NT_1\) gives a value for \(\tau_G\) of \(\approx 10^{-12}\) s and the n.O.e. is above 2.6 for \(\tau_G\) $10^{-12}$-$10^{-7}$ s. This model cannot fit the data for $\beta$ carbons of alanine in MBP.

Although the alanyl residues give rise to a single peak, which indicates homogeneous chemical shifts, the rotational reorientation rate about the Cx-C$\beta$ axis may not be the same for each residue. The motion of the main chain is also not isotropic, as a distribution of correlation times is needed to satisfy the relaxation parameters for the $\alpha$ carbons.

For phenylalanine side chains the angle of rotation (\(\theta\)) is taken as 60°. As with the alanyl methyl groups neither the isotropic model nor the single rotation model can satisfy the values of \(NT_1 = 133\) ms and n.O.e. = 2.2. For the isotropic model, \(\tau_R\) from \(NT_1\) is \(3.5 \times 10^{-10}\) or \(4.5 \times 10^{-8}\) s, and from n.O.e. it is \(1.5 \times 10^{-9}\) s. For the single internal rotation model with \(\tau_R\) 1 ns, \(NT_1\) gives a \(\tau_G\) of \(1.5 \times 10^{-9}\) s and the n.O.e. is above 2.6 for \(\tau_G\) $10^{-12}$-$10^{-7}$ s. Unlike the alanyl side chains the relaxation parameters of phenyl side chains can be fitted to a log\(\chi^2\) distribution of correlation times, \(p = 12\) (b = 1000) being used. A value of \(\bar{\tau}\) 0.2-0.3 ns fits the observed \(T_1\), n.O.e. and \(W_{\text{h/2}}\). As this mean correlation time defines a distribution of side-chain motions superimposed on a distribution of main-chain motions its physical significance is doubtful.

In conclusion, the dynamics of the side-chain carbons in this structured coil protein appear to be complex. Eventually studies based on enriched $^{13}$C residues may allow the dynamics of individual side-chain carbons to be elucidated experimentally. The combination of non-isotropic backbone motions with restricted rotation of the side chains leads to more parameters than can be evaluated from data on the average behaviour of amino acid residues.

**Structure and Function of Myelin Basic Protein**

In the course of evolution, the development of the vertebrate nervous system required the introduction of a material like myelin to permit rapid saltatory conduction in compact axonal processes. It has been suggested that MBP arose by modification of certain histone genes.\(^{42}\) We can surmise that MBP was an essential protein for the formation of the spiral insulating sheath of myelin. Several hypotheses have been advanced as to the precise role of MBP in the myelin structure. Two current models indicate that the function of MBP is to promote coherence between the pair of unit membranes of myelin at the major dense line or intracellular apposition. In one model, a single MBP molecule spans the intracellular cleft; in the other, a pair of molecules dimerizes across the cleft.

It can be suggested that structured regions of MBP associated with the PPP region may provide the extensions from the membrane surfaces that dimerize across the intracellular cleft. Further studies are in progress in our laboratory to obtain more

detailed information on the regions of the protein that are involved in self-association processes and in binding to the lipid moiety of the membrane.

Acknowledgments

This work is part of a program supported by the National Health and Medical Research Council, which is directed towards the problems of multiple sclerosis. We are thankful to Dr Ray Norton and (the late) Mr Leopold Klososzczyk for helpful advice and discussions, and to Dr Alan J. Jones for assistance at the National N.M.R. Centre. We thank Esso Australia Ltd for assistance in obtaining equipment.

Manuscript received 29 December 1977