NUCLEAR MAGNETIC RESONANCE STUDIES ON MYOGLOBIN

Thesis submitted for the degree of

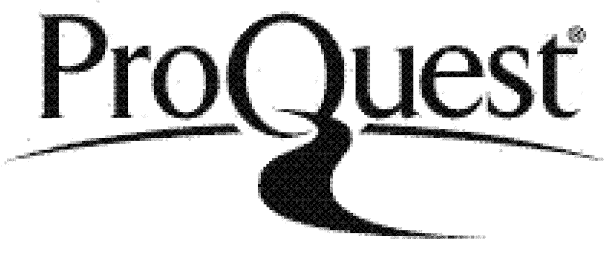
Doctor of Philosophy

at the University of Leicester

by

Christine Linda Wilkins

September 1992



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Nuclear Magnetic Resonance Studies on Myoglobin

Christine Linda Wilkins

Abstract

Nuclear magnetic resonance spectroscopy (NMR) studies of the paramagnetic haemoprotein equine cyanometmyoglobin are presented.

Many significant NMR assignments in both the proton and carbon-13 spectra of equine cyanometmyoglobin have been obtained. Notably the full proton and carbon-13 assignments for all four haem methyl groups have been obtained, resolving some previous ambiguity surrounding the assignment of the haem 3-methyl group, and full proton assignments for the haem 4-vinyl group, which is unresolved from the diamagnetic envelope in the one dimensional proton spectrum, have been obtained by use of two dimensional proton-proton shift correlation spectroscopy (COSY). Assignments are also presented for the ubiquitous proximal histidine residue, which plays a role in modulating the activity of the iron centre, as well as for certain other physiologically relevant haem pocket nuclei, providing a valuable basis for future studies of the structure-function relationships not only of equine cyanometmyoglobin, but also of those other related haemoproteins for which it can be used as a model.

It has been demonstrated that the COSY technique, which prior to the commencement of this work was believed to be inapplicable to paramagnetic species, since the line broadening and rapid loss of coherence caused by paramagnetic induced relaxation was thought to interfere with COSY peak detection, may indeed be profitably applied to the study of paramagnetic haemoproteins. The usefulness of other spectroscopic techniques, previously rarely used in the study of biological macromolecules, such as two dimensional heteronuclear shift correlation spectroscopy, and both one and two dimensional techniques involving the editing of the natural abundance carbon-13 spectrum according to the number of attached protons, to the study of paramagnetic haemoproteins has also been demonstrated.

To my Husband

DAVID

ACKNOWLEDGEMENTS

I wish to thank Dr. David Turner for his help, advice, and encouragement throughout the course of this work.

My thanks are also due to the Staff of the Department of Chemistry, and of the Biological N.M.R. Centre, University of Leicester, for their friendship and support, and particularly to Dr. Gerry Griffith, whose help and advice proved invaluable.

I am grateful to the Science and Engineering Research Council for their financial support.

Finally I would like to thank my husband, David, for his help with the word-processing, and for his constant support.

CONTENTS PAGE

CHAPTER 1

The Rationale for the Use of NMR Spectroscopy in the Study of Biological Macromolecules and an History of its Development

"To see a world in a grain of sand,
And a heaven in a wildflower,
Hold infinity in the palm of your hand,
And eternity in an hour."

William Blake
Auguries of Innocence

1.1 Aims and Objectives of this Research.

Nuclear Magnetic Resonance Spectroscopy (NMR) can be used to obtain valuable information on the structure and function of biological macromolecules in the physiologically relevant solution state, and may be considered to give information complementary to that obtained by means of X-ray diffraction techniques.

For the purposes of this research NMR studies on the paramagnetic haemoprotein equine skeletal cyanomet-myoglobin were undertaken.

Myoglobin is often used as the model molecule upon which new biophysical methods are tested, and NMR studies of haemoproteins are facilitated by the hyperfine shifts observed when paramagnetic forms of these proteins are employed. The nuclei which experience the largest hyperfine shifts, due to their proximity to the haem iron centre, are precisely those which are most intimately involved in the active site of the molecule, consequently their NMR behaviour may reveal key information on the protein structure, however this information may only be unlocked if unambiguous resonance assignments exist for those nuclei involved in the haem pocket, and such assignments are, therefore, the prime objective of this research.

It is further intended that this research should demonstrate the applicability of various spectroscopic techniques, rarely used in the study of biological macromolecules, and previously supposed to be ineffective in

the study of paramagnetic species, to paramagnetic biopolymers.

1.2 Advantages of NMR for the study of Proteins

In any consideration of the NMR spectroscopy of biological molecules in solution it is necessary first to address the rationale behind its use. NMR spectroscopic techniques now allow the spectroscopist to glean three dimensional structural information about proteins and other biological macromolecules, a field of information which was previously the sole preserve of X-ray crystallographers.

The determination of the three dimensional structures of these molecules by means of diffraction techniques utilising single crystals has an immediately obvious drawback, as illustrated by this quotation from a recent article on the genetic engineering of "new" proteins which appeared in the New York Times[1]; "....it can take several years to determine the structure of a single protein by X-ray crystallography. For one thing many proteins do not readily form crystals. Now nuclear magnetic resonance spectroscopy..... is being used to avoid crystallization altogether, because it can analyse proteins in solution."

Clearly, for those proteins for which no suitable crystals may be obtained, NMR studies are the only route to three dimensional structural information. It would be wrong however, to consider NMR as a convenient alternative to diffraction techniques or as a "fall back" option. Indeed Wüthrich[2] persuasively argues that the two techniques may be considered to be complementary to one another. The mere fact

that NMR spectroscopy is carried out in solution, whilst X-ray diffraction techniques employ single crystals, allows for direct comparison of protein structures in crystalline and non crystalline states, and indeed many such studies have been carried out. Typically the crystal and solution structures of protein molecules are found to be very similar as in the case, for example, of the study by Wagner et al[3] on the globular protein, basic pancreatic trypsin inhibitor, which found that apart from a few localised structural differences on the protein surface, the molecular architecture found in the crystalline state was largely conserved in aqueous solution. In addition to those proteins with reported similarities, however several startling structural differences between crystalline and solution state proteins have also been reported. Recent NMR studies by Shulze et al[4] on rat liver [Cd_7]metallothionein-2, for example, have shown that the solution structure of the protein is very similar to that of rabbit liver metallothionein, despite their primary structures differing by some 20%, although marked differences to the reported crystal structure[5] exist. The metal-cysteine coordinate bonds linking the 7 metal ions with the polypeptide chain are found to be different, and this, in turn, is believed to lead to the large observed differences in the polypeptide fold[4].

Similarly, NMR studies by La Mar and co-workers[6] have detected the presence, in various native myoglobins, of around 8-10% of a minor component in which the haem group is rotated

through 180° with respect to the protein matrix. La Mar suggests that the reason that these observations are not repeated by X-ray studies of single crystals might be that only the major form crystallizes[6].

Furthermore, since the biological functions of proteins invariably occur in solution it would seem logical that in any study of the structure function relationships of proteins the structure of relevance would be the solution structure. Indeed this advantage of solution studies may be further enhanced by matrix matching of the NMR sample to the biological environment, for example by adjusting the pH or ionic strength of the solution.

The facile adjustment of the experimental conditions in NMR studies leads to a further area of complementarity between the two techniques in that the effect on the protein of altering one or more of these conditions may be monitored.

The relationship between experimental temperature and hyperfine shifts in paramagnetic species is well documented, and will be discussed in more detail elsewhere in this work. The thermal variation of both carbon and proton resonances can yield a wide variety of information. For example temperature dependent conformational changes in proteins have been identified by means of mapping discontinuities observed in the temperature dependence of carbon-13 chemical shifts[7,8], and similar studies have shown that the binding of N-acetylglucosamine to lysozyme dramatically reduces the conformational flexibility of the protein[9]. The temperature

dependent behaviour of the amide protons in peptides has been shown to yield information about hydrogen bonding patterns[10].

Differential temperature dependencies of various chemical shifts can also play a role in the assignment of NMR spectra, particularly in crowded spectral regions, since resonances unresolved at one temperature may well be resolved at another.

Similarly coincident resonances may be differentiated by the variation of chemical shift position with pH. The pH dependence of chemical shifts has been employed to yield specific assignments[11]. NMR titration has also been used to illustrate the varying degrees of solvent accessibility of amino acids residues in globular proteins[12], and can yield information about the three dimensional structures[13].

Proteins cannot be considered to possess fixed rigid structures, and the two techniques of NMR and X-ray diffraction can also provide important complementary information due to the fact that the two types of measurement employ markedly different timescales, and therefore can yield dynamic information about the different types of fluctuations which can take place within molecules.

X-ray structure determinations provide an insight into those high frequency structural fluctuations with time constants of 10^{-11}s or less[14], which are mainly associated with the small rotational and vibrational motions seen in crystals. The X-ray structure, therefore gives an outline of the conformational space occupied by the molecule, allowing for these high frequency fluctuations, and inherently provides

roughly averaged atomic positions, about which the atoms themselves may exhibit a considerable degree of lower frequency motion.

For the characterization of those internal dynamics of protein molecules with time constants from approximately 10^{-9}s up to around one second NMR spectroscopy provides a powerful tool, since relaxation parameters bear a direct relationship to the amplitudes and correlation times of internal molecular motions. In common with X-ray diffraction techniques, NMR can provide information about the magnitude of the individual atomic motions, in contrast to X-ray diffraction however it is also sensitive to the timescales of those motions.

The dynamics within the NMR timescale range include many of those intimately involved with processes such as enzyme action and the structural rearrangements involved therein, and since one of the ultimate objectives of the study of protein dynamics is to relate the structure of a protein to its function, NMR is used extensively for this purpose. Roberts[15] for example, describes the use of NMR studies to characterise motions within the dehydrogenase enzyme, dihydrofolate reductase, and within various inhibitors bound to the enzyme, and discusses the importance of some of these fluctuations in the enzyme inhibitor interactions.

The fact that NMR can provide both structural and dynamical information in solution has also been employed by Dobson[16] to study the unfolded and partially folded states of several different proteins in order to derive information about

molecular events associated with the transitions undergone by a linear string of amino acid residues in the adoption of its final folded protein conformation. He hypothesises that this work may prove to be particularly valuable in those proteins whose functional states are partially unfolded, since it seems likely that crystallization of these states will prove to be impossible, and they are hence inappropriate for study by crystallographic techniques.

1.3 The Development of NMR Strategies for the Study of Proteins

Having addressed the rationale behind the use of NMR as a tool with which to study proteins it is interesting to consider its development as such a tool. The many applications of NMR to the study of biological macromolecules have evolved from the NMR spectroscopy of much smaller chemical compounds and natural products. The NMR phenomenon was first observed, independently in 1946 by both Felix Bloch[17] and Edward Purcell[18], but it was over 10 years later in 1957 that the first reported proton NMR spectrum of a protein appeared[19]. This report, by Saunders and co-workers[19], showed the spectrum of bovine pancreatic ribonuclease, and identified two broad spectral envelopes as aromatic and aliphatic protons. Given the available NMR equipment at that time it is unsurprising that the spectral resolution was low. Commercial spectrometers in the 1950's typically had proton resonance frequencies of 40MHz, improving to 60 or 100MHz in the 1960's. Further significant developments in this area of instrumentation did not come about until 1966, when the introduction of super conducting magnets allowed the

first high field NMR studies at a field strength of 200MHz[20], with a corresponding improvement in both the sensitivity and the spectral resolution of the NMR experiment. The next major step forward came in the early 1970's with the introduction of pulsed Fourier transform spectroscopy.

NMR spectra were originally recorded by scanning over the entire frequency range, and thus collecting a series of absorptions at successive frequencies, from which a spectrum might be constructed. This is known as the continuous wave (CW) spectrum.

If a high degree of resolution is to be maintained using CW NMR, relatively long sweep times are required as each individual frequency must be sequentially excited. It is however, possible to observed all the absorptive frequencies in a spectrum simultaneously, by effectively striking the sample with a single pulse of radiofrequency.

Heisensberg's uncertainty principle suggests that the shorter the duration of a nominally monochromatic pulse, the greater is the uncertainty about its frequency. Thus by applying a very short pulse (in the order of μs) the uncertainty is such that the whole frequency range of the spectrum may be stimulated.

The signal or interference pattern obtained subsequent to this pulse will contain all the characteristic frequencies of the sample, and will decay exponentially in the interval immediately following the pulse. The NMR signal in this form is known as the free induction decay (FID).

The FID is detected in a digital format, and reduced to the required frequencies by the subtraction of the carrier frequency, ν_o. The remaining signal is then converted using an analogue to digital converter.

Now the FID presents the signal as a function of time, and it is described as being in the time domain. A spectrum, however, displays the signal as a function of frequency, and is described as being in the frequency domain.

The time and frequency domains of a complex wave form are reciprocal to one another and can be interconverted by the Fourier transform (Eqn 1.1)

$$f(\omega) = \int_{-\infty}^{+\infty} f(t)\ e^{i\omega t}\ dt \qquad \text{Eqn. 1.1}$$

Where (t) is data in the time domain

(ω) is data in the frequency domain

The FIDs obtained by way of pulsed Fourier transform are additive, and consequently the pulse-signal-decay sequence may be repeated and added together many times, leading to an improvement in the signal to noise ratio (S/N), since true frequencies will interfere constructively, and random noise will interfere destructively in successive FIDs the S/N ratio improves in proportion to the square root of the number of scans (Eqn.1.2)

$$S/N \propto \text{scans}^{1/2} \qquad \text{Eqn. 1.2}$$

Early work on the NMR of proteins concentrated on the data that could be obtained from chemical shift measurements, since chemical shifts give direct information about the chemical environment of nuclei, and are very sensitive to conformational or chemical changes. Investigations of the positions of proton resonances of various amino acid residues in small peptides, and comparison of their positions to those random coil proteins have been carried out, and a high degree of correlation found. Such data have been employed in studies of protein denaturation[21,22]. Similarly the anomalous chemical shifts of protons affected by the ring current fields of local aromatic groups[21,23] was the subject of much early investigation. The sensitivity of chemical shifts to environmental changes was also investigated via the NMR pH titrations of ionizable groups[24].

The development of the use of Carbon-13 NMR to study proteins closely paralleled that of proton NMR. The first natural abundance Carbon-13 spectrum of a polypeptide; gramicidin SA, a decapeptide antibiotic, was published in 1970 by Gibbons et al[25] who noted that that carbon-13 NMR had many advantages over proton NMR for the study of polypeptides, in particular the large shifts and spectral widths, the relatively narrow line width, and the absence of spin-spin coupling, which lead to a simpler, more disperse spectrum, and also the near predictability of the chemical shifts, (see also Howarth and Lilley[26]).

Natural abundance carbon-13 spectra of proteins rapidly

followed. For example, Conti and Paci[27] compared the carbon-13 spectra of various globular proteins with the "composite spectra" of their constituent amino acids, and found that a high degree of correlation between the actual and the computed spectra existed, whilst Oldfield and Allerhand[28] observed the natural abundance carbon-13 spectrum of cytochrome-c, and differentiated the resonances of those non-protonated carbons near, or belonging to the central porphyrin group from others within the protein

Cozzone et al[7] used carbon-13 NMR to detect a temperature dependent conformational transition in lysozyme, and several workers[29,30] used the carbon-13 spectra of proteins, selectively enriched in ^{13}C at a particular amino acid residue, in order to make resonance assignments. At around the same time as the evolution of carbon-13 NMR as a tool with which to study proteins came two other important developments, namely the first reports of the observation of nuclear Overhauser effects in proteins[31,32], and the proposition by Jeener in 1971[33],of the idea of two dimensional NMR spectroscopy.

The Nuclear Overhauser Effect (nOe) is a relaxation effect which can occur in double irradiation experiments and arises from the dipolar coupling of nuclei in solution. The nOe is manifested as an increase or decrease in the intensity of the signal when a spatially neighbouring resonance is saturated.

As the nOe has its origin in dipolar relaxation, and the intensity of magnetic fluctuations at a nucleus is related to

the distance separating the mutually relaxing nuclei, its magnitude is dependent on the internuclear distance. The relationship between the nOe and internuclear distance is given by equation 1.3:-

$$I_{nOe} \propto r^{-6} \text{ ---------- eqn. 1.3}$$

where r is the internuclear distance. Consequently the nOe falls off rapidly with increasing r, and is normally only observable up to values of around r = 3.5Å.

The application of nOe measurements to proteins allowed the possibility of the NMR determination of inter proton distances in biological macromolecules, although initial work was hampered by spin diffusion effects, and it was not until 1978 that the nOe build up technique, which allowed estimations of inter proton distances in proteins, was introduced by Wüthrich and co-workers[34,35].

In 2D NMR experiments a second time variable is introduced such that a signal modulated as a function of a variable time parameter, t_1, is subsequently detected as a function of a second time variable, t_2. The data thus collected is, therefore, a function of two independent time variables, and just as Fourier transform converts a signal as a function of time into a signal as a function of frequency (Eqn. 1.4), so a two dimensional Fourier transform converts data from a function of two time variables into a function of two frequency variables (Eqn. 1.5), ie. a two dimensional frequency spectrum.

$f(t)$ 1D FT $f(\nu)$ ----------Eqn. 1.4

$f(t_1, t_2)$ 2D FT $f(\nu_1, \nu_2)$ ---------Eqn. 1.5

Jeener's experiment[33], which is also frequently referred to as COSY, gives rise to a two dimensional frequency spectrum in which both frequency axes are ^1H chemical shifts. Magnetisation components which have the same frequencies during both t_1 and t_2 appear on the diagonal of the two dimensional spectrum, which essentially shows the normal one dimensional proton spectrum. In a coupled system however, the application of a second pulse causes the magnetisation which arose from any one transition during t_1 to be transferred between all the other transitions to which it is coupled, and J coupling partners are therefore indicated by off diagonal cross peaks (peaks whose frequencies ν_1 and ν_2 are different), furthermore, since transfer of magnetisation between transitions may take place in both directions, a cross peak at coordinates (ν_1, ν_2) will possess a symmetrical partner at coordinates (ν_2, ν_1). This phenomenon can be used to distinguish true cross peaks from random noise, which will not display the same symmetry.

The advantages afforded by two dimensional techniques lie in the increased resolution resulting from the spread of data over a two dimensional frequency plane, rather than a one dimensional frequency axis such that the probability of accidental overlap is significantly lowered, and in the correlation data available.

It was not until 1977, however, following developments in

the theory and practice of two dimensional NMR by Ernst and co-workers[36,37], that two dimensional NMR methodology had advanced sufficiently to allow the first recording of a 2D NMR spectrum of a protein[38], namely the 2D, J resolved proton spectrum of the protein basic pancreatic trypsin inhibitor. These two developments have, in turn led to the development, by Wüthrich and co-workers, of an NMR technique, combining a procedure for the elucidation of sequence specific resonance assignments with the determination of inter proton distances within the molecule, for protein structure determination (the Wüthrich method of protein structure determination)[2,20], and the first full, proton NMR derived, solution conformation of a protein, proteinase inhibitor II from bull seminal plasma, was published in 1985[39].

1.4 An Analysis of Current Strategies for the NMR Study of Proteins

The pivotal NMR technique of the Wüthrich method of protein structure determination is two dimensional nuclear Overhauser spectroscopy (NOESY), complemented by the use of proton-proton shift correlated spectroscopy (COSY).

In the sequential assignment step of this procedure, developed by Wüthrich and co-workers in the late seventies and early eighties[40,41], the scalar spin-spin couplings observed in the COSY spectrum are used to unambiguously assign resonances as belonging to a particular type of amino acid residue. The NOESY spectrum is then used to establish the relationship

between sequentially neighbouring amino acids as NOESY cross peaks will be manifested by the close approach of the amide protons of neighbouring amino acids residues, and also the αCH of one amino acid residue and the amide proton of the next (Fig.1.i),

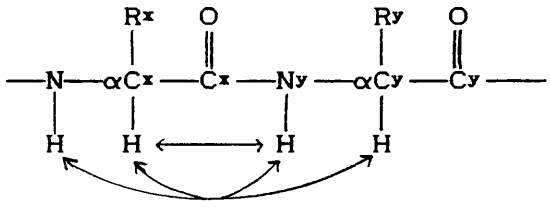


Fig.1.i An illustration of the short sequential distances between two neighbouring amino acid residues, x and y, which give rise to NOESY cross peaks.

Thus, since it is known from the data obtained from the COSY spectrum which type of amino acid residue x and residue y are, and from the NOESY spectrum that they are sequentially adjacent, this gives rise to the identification of a dipeptide which may be matched against the known protein sequence, and thus assigned. If the dipeptide segment is not unique within the amino acid sequence the method may be extended to identify tri or tetra peptides etc., until a unique fragment is found. In practice there is only a very small probability that a tri or tetra sequence will repeat within the total sequence of a globular protein[40].

Having established the assignments of resonances to the various amino acid residues, the second step is to establish

the secondary structure of the protein, and this three dimensional structural information is also available from the data obtained in NOESY spectra.

The presence of a NOESY cross peak shows that the two protons corresponding to the two correlated diagonal peaks are, in general, less than 5Å apart. The identification of NOESY cross peaks between non-sequential amino acids therefore gives valuable information about the three dimensional protein structure, showing the presence for example, of loops and folds within the polypeptide chain which allow the close approach of amino acids residues which are sequentially well separated. Complex patterns of nOe's can also arise which are characteristic of certain types of secondary structures. For example the inter proton distances involved between the protons of the first and fourth amino acid residues in an α-helix are around 4Å, compared with about 14Å in an extended polypeptide chain, consequently a repeated pattern of NOE's between pairs of amino acid residues these relative positions over a segment of the structure illustrates the presence of an α-helical section of polypeptide chain[20].

Since the NOESY spectra of proteins are very complex, and may contain hundreds of cross peaks, several mathematical algorithms have been developed to assist in the structural interpretation of this wealth of data. These algorithms are designed to transform the pairwise distance information obtainable from NOESY spectra into usable cartesian coordinates which allow the construction of valid structural models, and

rely on a library of constraints applied by the knowledge of the reasonable geometries of amino acid residues, to uncover a set of possible structures defined by both these constraints and the NMR data. An example is the distance geometry programme DISGEO[42,43], which is designed to identify those conformers where the set of pairwise distances between points fall between upper and lower limits defined by input nOe data and by a library of standard bond lengths, bond angles and van der Waals radii respectively. This programme was used in the first structure determination of a globular protein from experimental NMR data[39]. An analysis of this, and other similar programmes designed to generate three dimensional structures from NMR derived data is given in reference 44. A recent report on the NMR structure determination of the α-amylase inhibitor tendamistat[44a] contains a comparision of some of the available algorithms, as does the report on the structure of BPTI[3].

Although the Wüthrich method for protein structure determination has been found to be widely applicable in the case of proteins up to a molecular weight of 15,000, and has been used to solve in excess of 100 protein structures[2] it still has significant drawbacks. The presence or absence of cross peaks in a NOESY experiment is critically dependent upon the chosen value of the mixing time, during which the individual nOe's build up. If the mixing is too short cross peaks may be absent since the nOe's have not yet built up, if it is too long they may still be absent, the nOe's having died away. The nOe build up profiles of the various nuclei within a

sample will be different since they are dependent upon internuclear distances, therefore a NOESY experiment will only detect interactions between those nuclei whose internuclear distances are effectively selected by the chosen value of the mixing time. Consequently it is necessary to perform a set of NOESY experiments with various mixing times, and the non manifestation of a cross peak between two nuclei cannot be taken as evidence of the absence of any interaction between them.

In addition to this, with increasing protein size, an increase in the likelihood of spectral overlap, due to coincident, or nearly coincident, proton chemical shifts also occurs, which, coupled with the line broadening typical in the spectra of large proteins, can make the interpretation of the COSY fingerprint region, which is necessary for sequential resonance assignment, prohibitively complicated, and various methodologies have been proposed to deal with this problem. The simplest of these rely on the differing pH or temperature dependence of the individual proton chemical shifts, as described above.

A strategy proposed by Englander and Wand[45], and an algorithm based thereon[46], the main chain directed method for assignment of proton NMR spectra of proteins, relies on the identification of patterns of nOe connectivities, typical of secondary structural elements, within the NOESY spectrum, such that a more rigorous analysis of the amino

acid side chain system is not essential.

Isotope labelling strategies have been used to simplify the complex spectra of proteins, for example by labelling around 30% with carbon-13 and translating the data obtained from a series of two dimensional techniques; including $^{13}C-^{15}N$ and $^{13}C-^{13}C$ correlation experiments, to the fingerprint region of the COSY spectrum via a $^{1}H-^{13}C$ correlation, Markley and co-workers[47] have overcome many of the assignment ambiguities in the spectrum of ferredoxin from the photosynthetic cyano bacterium anabaena 7120. Similar strategies which employ the uniform labelling of proteins with nitrogen-15 have also been devised[48].

Various techniques involving deuteration of proteins have been successfully employed. Random fractional deuteration, in which all carbon bound protons are around 50-85% enriched with deuterium, simplifies homonuclear two dimensional proton NMR spectra, since the lower level of sample protonation decreases the incidence of proton-proton dipolar interactions. Since these are the dominant factors in proton relaxation times, this in turn leads to narrower linewidths[49,50].

A combination of random fractional deuteration and homonuclear two dimensional proton NMR has been utilised in the sequential assignment of thioredoxin from Escherichia coli[50]. Spectra may also be simplified either by the study of deuterated proteins, in which only a few amino acid types are protonated[51] or vice versa[50].

Three dimensional NMR techniques are also proving to be of value in the interpretation of protein spectra. Three dimensional techniques are an extension of two dimensional spectroscopy, by the addition of a further time period and are derived via "overlapping" combinations of two dimensional pulse sequences by means of omitting the detection time of the first experiment, and the preparation time of the second. The introduction of a third domain greatly increases spectral resolution as well as allowing some interesting correlations[52,53]. Homonuclear 3D experiments have been described, whereby, for example, a COSY experiment is extended into a third dimension to allow J modulation of the COSY cross peaks[54]. A homonuclear 3D NMR technique incorporating a Homonuclear Hartmann-Hahn pulse sequence was the first reported 3D technique to cover the full proton chemical shift range of a protein in all three domains, and was used in the assignment of pike parvalbumen[55].

Three dimensional techniques are also often used in conjunction with isotopic labelling. Heteronuclear three dimensional NMR, in which two dimensions are equivalent to a NOESY experiment whilst in the third dimension NH protons are labelled with their nitrogen-15 chemical shifts, has been used to obtain sequential backbone assignments in stapphylococcal nuclease, non selectively enriched in nitrogen-15[56].

Similarly a 3D combination of $^{1}H^{15}N$ correlation and

NOESY has been used to determine small heteronuclear coupling constants in nitrogen-15 labelled proteins[57].

The practicality of 3D NMR suffers from limitations due to the extremely long acquisition times involved, the three dimensional experiment on staphylococcal nuclease described above, for example, took 2½ days to run[56], and problems can also be associated with the storage and manipulation of the very large data sets involved, however it is likely that in the future 3D NMR will become a routinely used technique.

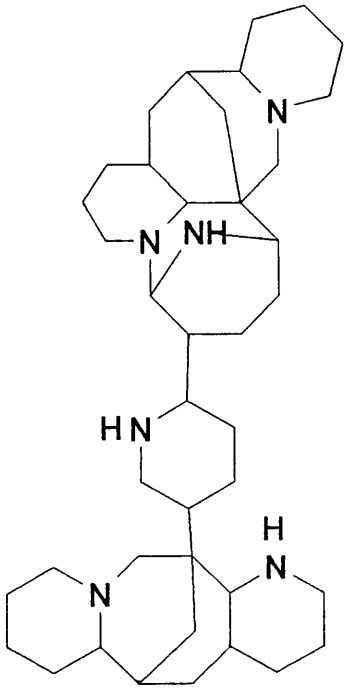
An application of 2D NMR spectroscopy which has so far found little favour in the study of protein molecules, except for comparatively recently when employing isotopically labelled samples[47,48], and when part of a 3D strategy[56,57] is that of heteronuclear shift correlation, the heteronuclear analogue of COSY. In heteronuclear shift correlation (HETCOR) experiments the two frequency axes represent the chemical shifts of two different nuclei, most often ^{1}H and ^{13}C, where a cross peak is manifested at the coordinates of a carbon nucleus and its attached protons. This procedure has several advantages, by correlation of proton chemical shifts with the much more disperse chemical shifts of the carbon spectrum, spectral overlap is largely removed.

In addition to this the carbon-13 chemical shift values may provide significant extra clues in the assignment of the proton spectrum, indeed proton resonances may be

cross assigned from known assignments in the carbon spectrum, and vice versa. Similarly the HETCOR experiment may be used in a complementary manner with the COSY spectrum, but although the first $^{1}H^{13}C$ correlation of a protein, carbon-13 enriched ferridoxin from anabaena variabilis, was published in 1982[58] with the types of cross assignments described above being reported for it the following year[59], and despite the technique having been successfully employed at natural abundance to obtain assignments for horse cytochrome c[60,61], this type of 2D NMR study of proteins is still rare.

This approach has however been widely used in the study of quite complex small natural products at natural abundance, for example in the study of bile acids[62] and to determine the previously unknown structures of the ormosia alkaloid ormisine[63] (fig 1.ii), and of 9,21-didehydroryanodine (fig 1.iii) a pesticide from the plant Ryania speciosa[64].

These and several other applications involving natural products are discussed by Derome[65]. Similarly a recent paper by Rinaldi and Swiecinski[66] discussed the application of COSY and a 2D experiment based on $^{1}H^{13}C$ heteronuclear Overhauser enhancement spectroscopy in the sequencing of small peptides.



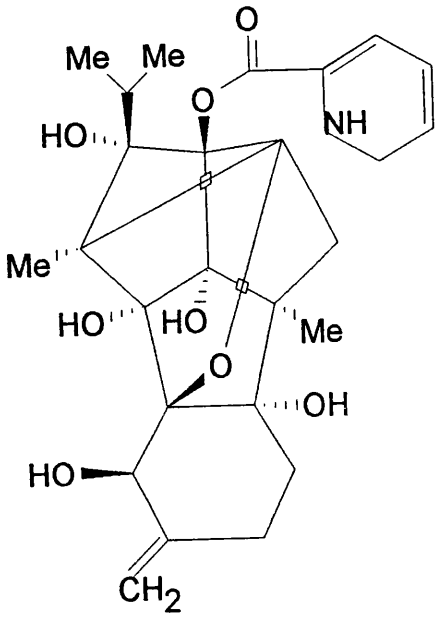


Fig 1.ii ormosinine Fig 1.iii 9,21-didehydroryanodine

Finally the role of less common nuclei in the NMR study of proteins cannot be ignored. Whilst biological macromolecules rarely contain fluorine-19 it may readily be incorporated into them by biosynthesis. The presence of ^{19}F has little effect on the biological functions of a molecule, but provides a convenient label whereby the molecules may be studied. Conformational changes in lysozyme, ribonuclease and haemoglobin have all been studied utilizing fluorine-19 NMR[67], and many other biological macromolecules, including bovine serum albumin[68], and α-chymotrypsin[69] have also been subject to ^{19}F NMR studies.

Many other nuclei are also of interest. The oxygen-17 chemical shifts, in aqueous solution, of the twenty common

amino acids have been published[70], and iron-57 NMR has been used in the study of cytochrome-c[71], whilst both proton and cadmium-113 NMR have recently been used to study the mode of cadmium binding to a "finger" peptide from an HIV-1 nuclei acid binding protein[72], providing a valuable model for the zinc binding which is believed to be necessary for the correct protein function of retroviral nuclei acid binding proteins[72].

CHAPTER 2

Haem Proteins and their Role in the Development of the NMR Spectroscopy of Biological Macromolecules

"Überhaupt ist es für den Forscher ein guter Morgensport, täglich vor dem Früstück eine Lieblingshypothese einzustampfen - das erhält jung."

Konrad Lorenz
Das sogennante Böse

2.1 The structures and functions of Haem proteins

Proteins containing metalloporphyrin groups are implicated in a wide variety of biological processes involving light absorption, electron transfer reactions, oxygen transport, and the breakdown of peroxides.

The primary light receptor in photosynthetic processes, chlorophyll, possesses a porphyrin group with magnesium as the centrally coordinating metal, whilst for haemoproteins with coordinately bound iron, a wide range of biological roles have been observed.

Iron protoporphyrin ix (haem), shown in figure 2i, provides the prosthetic group for the oxygen storage and transport proteins haemoglobin and myoglobin, whilst other iron porphyrin derivatives are prosthetic groups for the respiratory electron transport proteins, the cytochromes, involved in oxidative phosphorylation. The linkages between the porphyrin and protein chain also vary.

In the case of the electron transport protein cytochrome c the porphyrin group is attached to the polypeptide chain via two thioether linkages, whilst the fifth and sixth coordinate positions of the iron are occupied by the nitrogen of a histidine residue (His 18), and the sulphur atom of a methionine residue (Met 80). (Figure 2.ii).

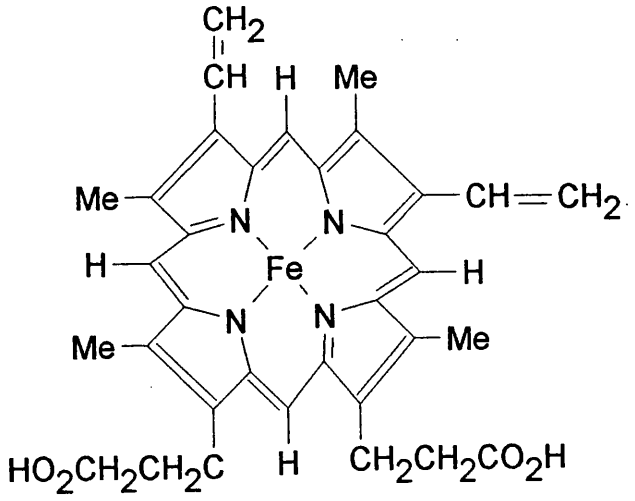


Fig 2.i

Iron Protoporphyrin IX

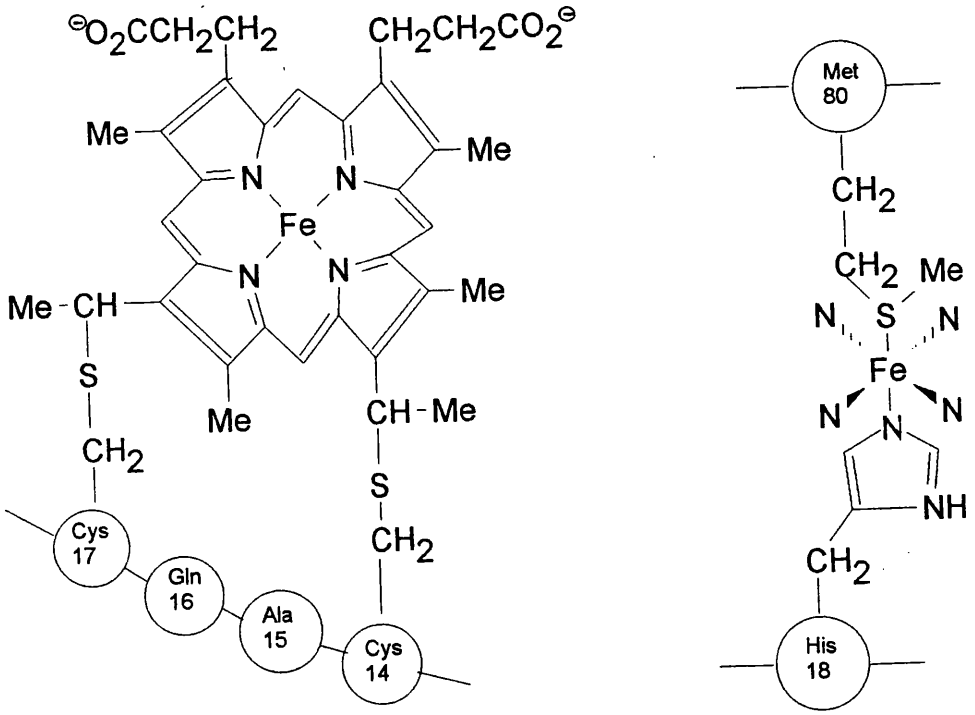


Fig2.ii

Attachment of haem group in cytochrome c

In contrast to this the haem moieties in myoglobin and haemoglobin are contained within haem pockets and are not covalently bound to the polypeptide chains, but are bound exclusively by coordinate linkage of the iron centre to the imidazole nitrogen of a histidine residue, known as the proximal histidine residue. This serves to weaken the ferrous oxygen bond in oxygenated derivatives of the proteins by donating electrons to the ferrous ion, and is thus responsible for the reversibility of oxygen binding in these proteins. The sixth coordinate position is occupied by a ligand which varies depending on the state and function of the protein. The attachment of the haem group in myoglobins is shown in Fig. 2.iii.

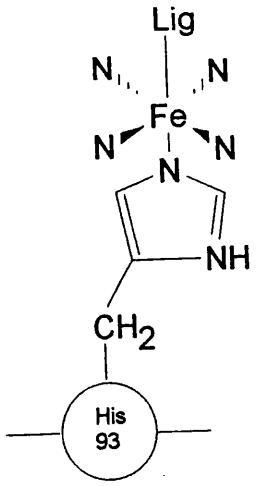


Fig 2.iii

Attachment of Haem Group in Myoglobins

Both ferrous and ferric, and both high and low spins states of iron are involved in the various biological roles of haemoproteins, and the electron configuration of the haem iron depends upon both its oxidation state and the nature of the ligand occupying the sixth coordinate position. Some examples are illustrated in Table 2.i.

The low spin ferrous state (FeII, S=0) is diamagnetic whilst the low spin ferric (FeIII, S=1/2), and the high spin ferrous (FeII, S=2), and ferric (FeIII, S=5/2) states are paramagnetic. The presence of a paramagnetic centre in these molecules has a profound effect upon their NMR spectra, and this will be discussed in greater depth in a later section.

Examples	Mb, Hb	MbO_2, HbO_2 Ferrocyt.c	$Mb.H_2O$ $Hb.H_2O$	MbCN, HbCN, Ferricyt.c
Oxidation state	Fe II	Fe II	Fe III	Fe III
Spin state	2	0	5/2	1/2
Electron configuration				
$d_{x^2-y^2}$	↑		↑	
d_{z^2}	↑		↑	
d_{xy}, d_{xz}, d_{yz},	↑ ↑ ↑↓	↑↓ ↑↓ ↑↓	↑ ↑ ↑	↑ ↑↓ ↑↓

Table2.i

Electron configurations of the Haem Iron

In the biologically active forms of myoglobin and haemoglobin the haem iron is in the ferrous state. When no ligand is attached the paramagnetic high spin state is observed, whilst the oxygenated forms are diamagnetic. If a cyanide ligand is attached at the sixth coordinate position then the complex is low-spin, and paramagnetic, but high-spin, and paramagnetic when this position is occupied by water. Both the oxidised and reduced forms of cytochrome c are low-spin.

Cytochrome c has a molecular weight of approximately 12,500, the polypeptide chain consisting of between 103 and 109 amino acid residues, depending on species. There is a single haem group per molecule, which is covalently linked to the polypeptide chain via two cysteinyl residues. The two axial positions of the haem are also occupied by amino acid residues, as illustrated in figure 2(ii).

Haemoglobin, which is responsible for the transport of oxygen from the lungs to the tissues, and carbon dioxide from the tissues to the lungs, via the bloodstream, consists of four sub-units of which two are known as α- and two as β-type sub-units. each of these sub-units comprises a polypeptide chain of around 150 amino acid residues, and an iron protoporphyrin IX moiety. The three dimensional structure of these individual sub-units takes the form of a characteristic "globin fold" of eight α-helical regions separated from one another by non-helical regions, (See Fig. 2 iv). The α-helices are designated

A to H, and the inter-helical junctional regions by a pair of letters representing the two helices which they separate. The haem group is located in a pocket between α-helices E and F.

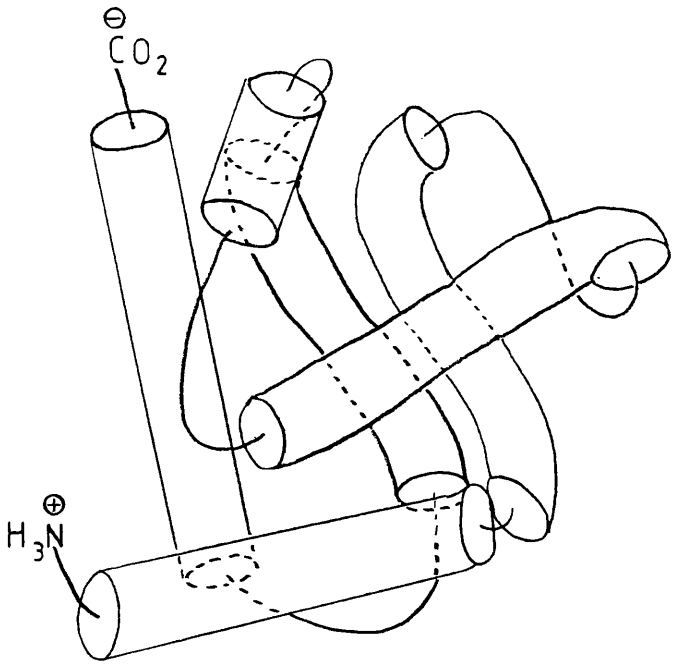


Fig 2 (iv)

The characteristic "globin fold" of oxygen binding haemoproteins, with the α-helical regions shown as cylinders.

Co-operative oxygenation, whereby partially oxygenated haemoglobin has a greater affinity for oxygen than does non-oxygenated haemoglobin, arises from a reversible transition of the haemoglobin molecule between two different structural states. The relaxed (R) state which has an oxy-haemoglobin like tertiary and quaternary structure, and a low-spin planar haem

iron, has a relatively high affinity for oxygen, compared with the taut (T) state, which has a high-spin deoxy-haemoglobin like structure in which the iron centre is approximately 0.07nm out of the plane of the porphyrin ring in the direction of the F helix.

Myoglobin is closely related to haemoglobin and is responsible for the binding and storage of oxygen in the muscle tissue. Myoglobin remains fully saturated with oxygen at much lower oxygen concentrations than does haemoglobin, only releasing it when the partial pressure of oxygen falls very low. Thus myoglobin fulfils its function of discharging oxygen,absorbed from the blood, to the mitochondria during periods of muscular exertion.

Myoglobin, which is responsible for the red colour of muscle tissue is particularly abundant in muscle tissue where oxygen shortage is likely to occur, for example the muscles of diving mammals such as whales and dolphins.

Myoglobins have a molecular weight of around 18,000, contain approximately 150 amino acid residues, depending on species, and consist of a single polypeptide chain and iron protoporphyrin IX group per molecule. Myoglobins take up the same characteristic "globin fold" as the individual sub-units of haemoglobin.

The amino acid sequences of many myoglobins are known[121]. Both sperm whale skeletal and equine skeletal myoglobin consist of 153 amino acid residues. The sequences of these two

myoglobins differ at some 18 points (17 substitutions, and 1 inversion), and a comparison of these two myoglobins, together with ovine myoglobin, is shown in table 2.ii.

Oxygen binding haemoproteins contain only two amino acid residues which are totally constant, regardless of species[120], namely the proximal histidine residue (His F8), and the phenylalanine (Phe CD1) residue located at the CD helical junction. They are , nonetheless, extremely sensitive to amino acid sequence, with severe functional disorders such as sickle cell anaemia being caused by a single amino acid substitution.

Table 2.ii

Amino Acid Sequences of Three Myoglobins

Sperm Whale:	Val-Leu-Ser-Glu-Gly-Glu-Trp-Gln-Leu-Val	10
Equine:	Gly Asp Gln	
Ovine:	Gly Asp	
Sperm Whale:	Leu-His-Val-Trp-Ala-Lys-Val-Glu-Ala-Asp	20
Equine:	Asn Gly	
Ovine:	Asn-Ala Gly	
Sperm Whale:	Val-Ala-Gly-His-Gly-Gln-Asp-Ile-Leu-Ile	30
Equine:	Ile Glu-Val	
Ovine:	Glu-Val	
Sperm Whale:	Arg-Leu-Phe-Lys-Ser-His-Pro-Glu-Thr-Leu	40
Equine:	Thr-Gly	
Ovine:	Thr-Gly	
Sperm Whale:	Glu-Lys-Phe-Asp-Arg-Phe-Lys-His-Leu-Lys	50
Equine:	Lys	
Ovine:	Lys	
Sperm Whale:	Thr-Glu-Ala-Glu-Met-Lys-Ala-Ser-Glu-Asp	60
Equine:		
Ovine:		

```
Sperm Whale: Leu-Lys-Lys-His-Glu-Val-Thr-Val-Leu-Thr      70
Equine:                          Thr-Val
Ovine:                           Asn-Val

Sperm Whale: Ala-Leu-Gly-Ala-Ile-Leu-Lys-Lys-Lys-Gly      80
Equine:                 Gly
Ovine:                  Gly

Sperm Whale: His-His-Glu-Ala-Glu-Leu-Lys-Pro-Leu-Ala      90
Equine:
Ovine:                           Val      His

Sperm Whale: Gln-Ser-His-Ala-Thr-Lys-His-Lys-Ile-Pro      100
Equine:
Ovine:       Glu              Asn

Sperm Whale: Ile-Lys-Tyr-Leu-Glu-Phe-Ile-Ser-Glu-Ala      110
Equine:                                        Asp
Ovine:       Val                               Asp

Sperm Whale: Ile-Ile-His-Val-Leu-His-Ser-Arg-His-Pro      120
Equine:                                  Lys
Ovine:                              Ala-Lys

Sperm Whale: Gly-Asn-Phe-Gly-Ala-Asp-Ala-Gln-Gly-Ala      130
Equine:
Ovine:       Ser

Sperm Whale: Met-Asn-Lys-Ala-Leu-Glu-Leu-Phe-Arg-Lys      140
Equine:          Thr                               Asn
Ovine:           Ser                               Asn

Sperm Whale: Asp-Ile-Ala-Ala-Lys-Tyr-Lys-Glu-Leu-Gly      150
Equine:
Ovine:           Met         Glu          Val

Sperm Whale: Tyr-Gln-Gly                                  153
Equine:      Phe
Ovine:       Phe
```

2.2 The Particular Role of Haemoproteins in the development of NMR Spectroscopy of Proteins.

Haemoproteins have played an important role in the development of NMR spectroscopy of proteins, not only because they are physiologically interesting molecules, but also because of the attractive spectral properties which they possess.

The large local magnetic field of the porphyrin group present in haemoproteins gives rise to ring current shifts very much higher than those cause by the aromatic rings of histidine, phenylalanine, tryptophan and tyrosine residues[73] as a result of which several resonances are well resolved from the main protein envelope.

Additionally haemoproteins can exist in paramagnetic forms in which the haem iron functions as an inbuilt paramagnetic shift probe, the unpaired electron causing large hyperfine shifts, which affect the resonances of those nuclei in close approach to the haem iron, allowing resolution of peaks for nuclei at, or near, the active site.

The existence of well resolved resonances in the spectra of both diamagnetic and paramagnetic haemoproteins permitted certain NMR experiments to be carried out on them prior to the advent of the two dimensional techniques which would otherwise have been required to give adequate levels of resolution. Consequently many types of NMR studies were first carried out on haemoproteins. The first observation of proton-proton nOe's

in a protein, for example, were made with cytochrome c[31], and Wüthrich[20] describes haemoproteins as important stepping stones in the development of his sequential resonance assignment procedure, and subsequent structural interpretation of NMR data.

Some haemoproteins also have a distinct advantage for NMR studies due to the facile extraction and subsequent re-insertion of the haem prosthetic group. This allows both the study of compounds in which the haem group has undergone some modification such as isotopic labelling, as in, for example, the study carried out by Mayer et al[74] to identify the proton resonances of three out of the four haem methyl groups, and the comparison of the spectra of entire proteins with those of the equivalent proteins, from which the prosthetic group has been removed, and of lone prosthetic groups[75].

The popularity of haemoproteins as the subjects of NMR studies is well illustrated by the abundance of available literature on this topic. A search of Chemical Abstracts files revealed 183 publications on the NMR of myoglobin alone, and it is on the NMR of myoglobin that the remainder of this review will focus.

2.3 A Review of NMR Studies on Myoglobin.

Any detailed structural or functional interpretation of the NMR data obtained from proteins relies on the availability of unambiguous assignments for the pertinent resonances.

Since the early studies carried out by Wüthrich[75], proton

and/or carbon assignments have been reported for parts of 26 out of 153 amino acid residues of sperm whale myoglobin and for certain of the haem resonances, and these are summarised in Table 2.iii

Early NMR studies on myoglobin involved the straight forward observation and comparison of different spectra, and concentrated mainly on the identification and assignment of those resonances which experienced large ring current shifts caused by the presence of the haem, or in paramagnetic haem compounds those affected by hyperfine shifts[76]

In 1968 Bak et al[23] observed that certain proton resonances in the spectra of both myoglobin and haemoglobin, were differently affected in the oxy and deoxy forms of these proteins, whilst Wüthrich[75] in 1970 published partially assigned spectra of isolated porphyrins, which were subsequently utilized in studies of the reconstituted protein, using both the native porphyrin, and deuteroporphyrin derivatives[76,77], and early comparison of the positions of some of these highly shifted resonances in the spectra of both oxymyoglobin and deoxymyoglobin[76], were used to identify the occurrence of conformational changes on the oxygenation of myoglobin and to estimate the magnitude of the displacement experienced by phenylalanine residue 43 during those changes.

Others workers also studied the spectral changes which occurred upon the binding of myoglobins, and other haemoproteins, with species other than oxygen, including

Table 2.iii Assignments available in the Literature.

2.iiia Proton

Assignment	Chem. shift	pH	Temp.	Rationale	Ref.
Haem 5 Me	27.4				
Haem 1 Me	18.7	6.8	298K	selective	74
Haem 8 Me	13.0			deuteration	
Haem 3 Me	-1.7	9.0	295K	from $^{1}H^{13}C$ COSY following ^{13}C assignment	134
Haem 3 Me	5.1	9.0	308K	nOe difference, deuterated haem	108
Haem 2 vinyl CαH	17.5			nOe difference	135
Haem 2 vinyl CαH	17.9	8.0	298K		87
Haem 2 vinyl CβH trans	-2.44				
Haem 2 vinyl CβH cis	-1.78				
Haem 2 vinyl CβH trans	-2.6	10.4	298K	nOe difference isotope labels	136
Haem 2 vinyl CβH cis	-1.8				
His 64 CδH (distal)	11.6			nOe difference	135
His 93 CβH (proximal)	11.5			nOe difference	135
Phe 43 p-H	12.7	8.6	308K	nOe difference	120
Phe 43 m-H's (av)	16.7				
His 64 NH	22.0	8.56	313K	Calculation of internuclear distances from spin lattice relaxation times	118
His 93 NH	20.0				
His 93 peptide H	13.5				
Tyr 146 H2,6	5.81				

Assignment	Chem. Shift	pH	Temp.	Rationale	Ref.
Tyr 146 H3,5	6.30				
His 24 H4	6.45				
Tyr 103 H2,6	6.45				
Tyr 103 H3,5	6.66				
His 97 H2	6.76				
His 113 H4	6.8				
His 36 H4	6.95				
His 48 H4	6.95				
His 12 H4	7.04	8.4	293K	Temperature and pH dependencies of shifts	13
His 119 H4	7.17				
His 116 H4	7.20				
His 32 H2	7.42				
His 24 H2	7.72				
His 113 H2	7.73				
His 116 H2	7.73				
His 48 H2	7.76				
His 12 H2	7.78				
His 81 H2	7.84				
His 81 H4	7.86				
His 119 H2	8.33				
Ile 99 $C\gamma H_3$ & $C\gamma_1 H$	-2.3				
Ile 99 $C\delta H_3$	-2.5	8.6	Various	nOe difference	136
Ile 99 $C\gamma H$	-7.8				
Ile 99 $C\alpha H$	-0.7				
Ile 99 $C\beta H$	-0.3				

Table 2.iiib Carbon

Assignment	Chem. Shift	pH	Temp.	Rationale	Ref
Haem 5 Me	-60.5			From $^{1}H^{13}C$ COSY following ^{1}H assignments in Ref 13	134
Haem 1 Me	-39.3	7.5	295K		
Haem 8 Me	-33.2				
Haem 3 Me	-28.5	7.5	295K	Due to magnitude of upfield shift and similarity of width to other Haem methyls.	134
Haem 2 vinyl Cα	51.3			By substitution of a ^{13}C labelled haem into the protein	137
Haem 4 vinyl Cα	81.2	8.6	298K		
Haem 2 vinyl Cβ	186.7				
Haem 4 vinyl Cβ	171.9				
Ile 28 $C\delta H_3$	14.16			Identified as Ile $C\delta H_3$s by position in "unique" range of spectrum, and differentiated by pH dependency and nOe's.	115
Ile 30 $C\delta H_3$	13.17				
Ile 112 $C\delta H_3$	12.85				
Ile 111 $C\delta H_3$	12.51				
Ile 107 $C\delta H_3$	12.75				
Ile 142 $C\delta H_3$	11.40				
Ile 101 $C\delta H_3$	9.80				
Ile 75 $C\delta_{H3}$	9.20				
Gly 153 C=O C	176.2	8.64	312K	By NMR titration	138
Met 55 $C\varepsilon H_3$	17.8	7.3	ambient	By ^{13}C labelling of Met @ $C\varepsilon H_3$	139
Met 131 $C\varepsilon H_3$	15.9				

chlorine binding to carbonmonoxymyoglobin[78], and formate[79], and xylidine[80] binding to myoglobin. Similarly early Carbon-13 NMR studies[27] involved the comparison of the spectra of myoglobin, and other haemoproteins both with one another, and with composite spectra of their constituent amino acids.

Due to the more disperse nature of carbon NMR spectra certain aromatic amino acid residue carbons give rise to well resolved resonances in an otherwise uncrowded region of the spectrum, and consequently these were among the first carbon resonances of myoglobin to be identified, studied and assigned[11,81].

It soon became clear that NMR spectroscopy provided a powerful probe of haem structure, and environment, and could be used to gain a unique insight into various aspects of the structure to function relationships of haemoproteins, providing information, for example, on haem-protein interactions or on the intermolecular motions and rearrangements connected with protein function.

The NMR investigation of the structure of haemoproteins, and the relationships of such structures to protein function is greatly simplified in paramagnetic forms of the proteins due to the presence of large hyperfine shifts which permit the resolution of peaks for nuclei at or near the active site, and many of the studies undertaken on myoglobin utilise the low spin, paramagnetic cyanide complex of myoglobin (cyano-metmyoglobin, MbCN), which is isostructural with the physiologically important carbonmonoxy derivative, or other

paramagnetic myoglobin complexes.

Since various haemoproteins with identical ligand binding sites and essentially very similar arrangements of their polypeptide chains can show marked differences in their ligand affinities, and in their haem ligand reaction rates, it can be inferred that the oxygen binding and redox properties of the central metal in haemoproteins are intimately controlled by local effects due to groups in close approach to the haem[82]. It is precisely these near haem-groups whose NMR signals will be affected by hyperfine shifts in paramagnetic haemoproteins and the use of hyperfine shifts as a probe of the environment of the prosthetic group in haemoproteins was postulated by Frye and La Mar in 1974[83].

Hyperfine shift patterns have subsequently been shown to be highly sensitive to the small local differences found between closely related genetic variants of myoglobin[84,85], and to those created in synthetically derived point mutants of myoglobin[85], and undoubtedly provide valuable probes into this modulation of haem activity by the surrounding amino acid residue pocket.

Hyperfine shift patterns are also extremely sensitive to the nature of the coordinated ligand. Iizuka and Morishima[86] showed that for myoglobin complexes the magnitude of the observed hyperfine shifts, as well as the signal width, increased in the order cyanide > azide > imidazole > deuteroxide, and observed anomalous temperature dependencies of the hyperfine shifted haem methyl signals of the azide and

imidazole complexes, compared to those of other paramagnetic myoglobins, with the two higher field signals shifting to lower field, and the two lower field signals shifting to higher field with increasing temperature.

A later study by La Mar and co-workers[87], using myoglobins substituted with selectively deuterated haems demonstrated that two different patterns for haem methyl hyperfine shifts are observed, dependent on whether the complex is wholly, or predominantly, high or low spin. The anomalous variable temperature behaviour of the azide and imidazole complexes of myoglobin observed by Iizuka and Morishima[86] are consequently explained as these complexes undergo a transition from low to high spin state with increasing temperature, so that the NMR signals are observed crossing over from the low spin to the high spin pattern.

The hyperfine shift pattern of the haem methyls in these haemoproteins appear to be directly correlated with the dominant spin state of the ferric iron centre, and the larger effect in low spin complexes is suggested to be due to the transferred π spin density being more susceptible to perturbations than the relatively insensitive σ system involved in the dominant spin transfer in high spin complexes[87].

Indeed studies of hyperfine shift patterns have shown that they are sensitive not only to the spin state of the central iron but also to its oxidation state, and to any ligand present[88,89,90]. Such hyperfine shift pattern studies have been used to characterise the iron porphyrin complexes in

haemoproteins by comparison with model compounds. Proton NMR data has been used, for example, to demonstrate the presence of a low spin ferryl oxo porphyrin ($Fe^{IV}=O$) both in hydrogen peroxide treated myoglobin, and in compound II of horseradish peroxidase, an intermediate structure present during the reaction of horseradish peroxidase[88], and detailed NMR studies of synthetic ferryl myoglobin have subsequently been undertaken to more closely characterise the active site in this complex[91] which serves as a useful model, of the active site structure and mechanism of peroxidase activation, and of the suggested active oxidising form of the enzyme cytochrome P450.

Similarly carbon-13 NMR of various ligated states of myoglobin and model compounds containing a haem group, specifically enriched in carbon-13 at the β carbons of the 2 and 4 position vinyl groups has shown that the chemical shifts of these two hyperfine shifted resonances are sensitive to both the spin state of the central iron, and consequently to ligand binding, and to the immediate protein environment of the haem[90].

The observed differences in the variable pH behaviour of the distal histidine (His 64) residue H2 resonance on going from sperm whale oxy myoglobin to sperm whale carbonyl myoglobin[92] provides a further example of the sensitivity of NMR parameters to the functional state of the haem group. These differences have been postulated to arise due to a shift away from the arginine 45 residue, as well as possibly from the iron itself (although this would have to occur such that there was

no appreciable alteration in the magnitude of the porphyrin ring current effect experienced) of the imidazole ring on going from carbon monoxide ligated to oxygen ligated myoglobin[92]. This movement may be possibly be involved in some "gate like" function of the distal histidine residue in ligand binding.

Since it has been shown[93] that there is a relationship between ligand binding to haem prosthetic groups and the tertiary structure of haem proteins this type of study is obviously of interest, and many workers have employed isotopically labelled ligands as probes of the haem-ligand contact environment.

Early studies using a cyanide ligand enriched in carbon-13 as a potential probe of ferrihaemoprotein structure were carried by Goff[94], whilst Morishima and Inubushi carried out nitrogen-15 NMR studies of the binding of cyanide, labelled with nitrogen-15, to myoglobins, cytochromes and haemoglobins[95], and studied both solvent[96], and pH[97] effects on the paramagnetic shift of the labelled ligand, and more recent studies have compared these shifts with their equivalents in nitrogen-15 labelled cyano peroxidase complexes[98].

The sensitivity to environmental changes of the chemical shifts of resonances other than those arising from the porphyrin or its bound ligands may also be used to glean information about protein structure and function.

A study of the titration behaviour of the chemical shifts of the C_γ and C_ζ of individual tyrosine residues in various myoglobins[12] has been used to yield information about the

environments of these residues in solution, showing that for sperm whale myoglobin one tyrosine residue (Tyr 151) is relatively exposed to solvent, one (Tyr 146) is buried in the interior of the protein molecule, inaccessible to solvent, whilst the third (Tyr 103) shows intermediate behaviour showing partial solvent exposure. These conclusions suggest that the solution structure of sperm whale myoglobin is consistent with its crystal structure with respect to the environments of the three tyrosine residues.

Similar studies have demonstrated the levels of accessibility to solvent of certain histidine residues in sperm whale oxy- and carbonmonoxymyoglobin[92], and the titration behaviour of the H2 resonance of histidine residue 36 has been demonstrated to be different in paramagnetic myoglobin complexes to that observed in diamagnetic myoglobins, due to the presence at low pH in paramagnetic complexes of a hydrogen bond to the glutamine 38 residue, which is broken at high pH[13].

Studies of the pH dependence of the methyl carbons of methionine residues of sperm whale myoglobin, specifically enriched with carbon-13 at those carbons, have been used to assess the mobility of the side chains of the methionine residues, and have shown that their chemical shifts are ligand dependent, despite being approximately 17Å removed from the iron centre[29].

It has been shown[99] that the proton chemical shift of the haem 1-methyl group of sperm whale metmyoglobin cyanide does

not display pH dependence, whilst those of haem 5-methyl, and haem 8-methyl groups do.

NB. The proton resonance of the haem 3-methyl group is not resolved from the diamagnetic envelope. This has been explained in terms of the relative isolation of the haem 1-methyl from the proximal and distal histidine residues compared with the haem 5- and 8-methyls[99].

NMR studies have also been employed to reveal structural changes experienced under high pressure by horse cyanometmyoglobin[100,101,102]. These changes appear to be restricted to the distal side of the haem pocket which experiences localised compressibility, for example, the distal histidine is thought to experience a dislocation away from the iron centre, whilst the proximal side remains unaltered, as evidenced by the chemical shift of the distal histidine residue N$_{\varepsilon}$H resonance, which experiences a sizable change in chemical shift on going to high pressure, whilst that of proximal histidine residue is unaffected[101]. The fact that the chemical shifts of the haem 8- and 5-methyls experience an upfield shift at high pressure, whilst that of the haem 1-methyl resonance does not, together with similar observations on the chemical shifts of the haem vinyl resonances further support the localised compressibility of the haem distal side[100]. Much investigative work has been undertaken into the precise nature of the relationship and interaction between a porphyrin and its associated polypeptide chain.

It has long been known that a certain degree of

structural heterogeneity, with respect to the orientation of the haem group within its pocket, exists in solutions of haemoproteins. La Mar and co-workers observed, by proton NMR approximately 10% of a secondary form of sperm whale metmyoglobin, in 1983[6], and similar observations have subsequently been made in the proton NMR spectra both of other myoglobins, and of haemoglobins[103,104,105]. By reconstituting apomyoglobin with haems isotopically enriched at each individual haem methyl site and observing the resultant positions of the characteristic haem methyl resonances it has been shown that the two forms of myoglobin differ in that the haem is rotated within its pocket by 180° about its α-γ meso axis when going from the major to the minor form. This structural heterogeneity, which is not found in single crystals[106], suggesting the possibility that only one form crystallises, is also observed in various other physiologically important forms of the protein, leading to the conclusion that the observed physiological properties of myoglobin must, in fact, represent a weighted average of the properties of its major and minor forms. Furthermore, it has been shown[6,107], that the reaction between apomyoglobin and haem initially yields a 1:1 mixture of myoglobins with the two different haem orientations, which slowly equilibrates to give a mixture in the observed native proportions.

Clearly since a 1:1 mixture of the two isomeric complexes is initially formed, the mechanism of the haem apoprotein reconstitution cannot differentiate between the two alternate

haem orientations. A possible mechanism which could lead to this lack of orientational recognition would be one in which the initial haem apoprotein interaction is the formation of either or both of the salt bridges between the propionate groups and their amino acid contacts in the haem pocket. NMR studies of the ratios of initial complexes formed on the reaction of apomyoglobin with, in turn, each of the pair of porphyrins in which one of the haem propionate groups has been selectively replaced by a methyl group[108,109], have shown that the haem propionate groups are indeed involved in the first step in the insertion mechanism. Furthermore each of the modified porphyrins shows a strong preference for the initial orientation with respect to apomyoglobin which places its single propionate group in the position which is characteristic of that of the 6-propionate group in single crystals. Crystal studies have shown that the haem 6-propionate group forms a salt bridge with the amino acid residue arginine 45[106,109], so it is likely that the mechanism of haem insertion into myoglobin involves an initial interaction between apoprotein residue Arg. 45 and one of the two haem propionates. It is further believed that due to the similarity of the half lives of the minor forms of sperm whale cyano metmyoglobin regardless of whether native haem or either of the two mono-propionate haems is incorporated[109,110], that salt bridge cleavage is not rate determining in haem reorientation and therefore the reorientation of the haem is likely to involve its spontaneous release from, followed by its reincorporation into the haem

pocket.

Proton NMR studies of cyano metmyoglobins reconstituted with various modified meso-tetraalkyl haems have shown that these synthetic porphyrins are mobile, and rotate freely within the haem pocket. It is suggested that this free rotation is consequent upon the disruption of the highly stereospecific haem-globin contacts, and occurs via a series of small fluctuations in the haem-apoprotein interactions, and that these fluctuations also occur in the native protein in which similar rotation is hindered by the close packing of the haem cavity[111,112].

NMR spectra of the major and minor forms of the sperm whale cyano metmyoglobin have been used as comparative models in the study of other, less well characterised haemoproteins, for example, a study comparing the NMR spectrum of sperm whale metmyoglobin with that of haemoglobin from the marine annelid, Glycera dibranchiata[105], has shown that this haemoprotein also exists as an isomeric mixture of major and minor forms. The assignment and comparison of the hyperfine shifts of the haem methyl groups has shown that the major form of sperm whale myoglobin is equivalent to the minor form of Glycera dibranchiata haemoglobin, and vice versa. Additionally the fact that the equivalent pairs of sperm whale myoglobin, and Glycera dibranchiata haemoglobin exhibit identical patterns of haem methyl hyperfine shifts has been taken as evidence of essentially very similar haem environments in the two proteins, despite the unusual distal leucine residue of Glycera

dibranchiata haemoglobin.

The much larger spread of haem methyl hyperfine shifts in various paramagnetic haemoproteins compared to that observed in the related model compounds is indicative of some degree of asymmetry in the haem environment, which is presumably induced by the haem apoprotein interactions. These interactions perturb the axial symmetry of the haem, shifting the major magnetic axis from its expected position directly along the haem normal.

NMR studies of the low spin iron III cyano-met. forms of sperm whale myoglobin, elephant myoglobin (in which the distal histidine residue is replaced by a glutamine residue), and various distal point mutants of myoglobin have been undertaken to characterise this process[86], concluding that observed differences in the hyperfine shift patterns of these cyano metmyoglobins appear to bear a direct relationship to the tilt of the major magnetic axis from its expected haem normal position, and in turn that this tilt in the magnetic axis is indicative of the steric tilt induced in the Fe-CN moeity by haem protein interactions.

This protein induced tilt in the cyanide ligand is of interest since cyanide and carbonmonoxy ligands are isostructural, and therefore it may be used as a model of the distortion induced in haemoprotein Fe-CO units. This distortion is believed to be responsible for the discrimination, by the haem binding site in intact myoglobins and haemoglobins, between oxygen and carbon monoxide, a relative instability being induced in the linear geometry of the carbon monoxide

molecule, compared with the more sterically stable bent geometry of the oxygen molecule[85,113].

It has long been known that the physiological functions of proteins are not merely functions of their static crystal structure, but are modulated by dynamic fluctuations of the proteins molecules. Indeed in the case of myoglobin such fluctuations are of particular interest as they are believed to permit the passage of oxygen between the protein exterior and the haem binding site, this ligand entry channel through the protein matrix being apparently locked in static crystal structures[93,106].

Individual small fluctuations within proteins which may be monitored by NMR include low to intermediate frequency motions ($0-10^5 sec^{-1}$) such as the "ring flips" of aromatic amino acid residue side chains, which may be observed by chemical exchange phenomena, and higher frequency motions ($10^8-10^{12} sec^{-1}$) such as those which occur in aliphatic amino acid residue side chains, which can be derived from nuclear relaxation processes[114,115].

It is believed that the internal motions within globular proteins are correlated, that is that groups of atoms from different amino acid residues move together, causing the protein to switch from one equilibrium structure to another, and that these types of local rearrangements must be responsible for the opening of a ligand entry channel in myoglobins. Evidence for this correlated motion theory has been obtained from carbon-13 NMR data from studies of the motions of

the isoleucine residues in sperm whale cyano metmyoglobin[115,116].

The sperm whale myoglobin molecule is a closely packed structure with seven out of the nine isoleucine residues therein being located in relatively constrained helix-helix contact regions[117], however comparison of the motions of these isoleucine residues with their environments has shown that the greatest motional amplitudes are associated with precisely those residues which are located in the most densely packed regions of the molecule[116], such that some complimentary motion of their neighbouring structures is implied, and the theory of correlated motion supported. Studies of the dynamics of labile hydrogen exchange in the haem cavity have also been employed to probe the potential ligand entry channels of myoglobin[118,119], and the exchange rate profiles of protons bound to haem cavity histidyl nitrogen atoms in sperm whale, equine, and canine cyano metmyoglobin have been shown to favour a penetration model of ligand entry[119], whereby the ligand passes through the protein matrix via an interconnecting network of small, discrete, transient cavities, over a local unfolding model.

Similarly the dynamic characterisation, by proton NMR, of the ring flipping experienced by the phenylalanine 43 residue in sperm whale cyano metmyoglobin has been undertaken[120]. This has shown that the activation energy required for phenyl ring rotation in this important amino acid residue; which is one of only two completely invariant amino acid residues in haem oxygen binding proteins[121]; and which is in Van der Waals

contact with the distal side of the haem π system[106]; is consistent with local concerted structural fluctuations of the CD helical corner at which it is located, again favouring a penetration model of ligand entry.

In addition to proton and carbon-13 NMR studies of myoglobins various interesting investigations utilizing less usual nuclei have also been undertaken.

Both myoglobins and haemoglobins enriched in Iron-57 have been studied by Iron-57 NMR[122,123], and their relaxation parameters determined, whilst carbon-13, Iron-57 double enriched myoglobin has been the subject of carbon-13-Iron-57 double resonance experiments[124,125], with the intention of characterising changes in the haem iron environment in the various functional forms of different haemoproteins.

Porphyrin has been substituted with trifluoromethyl groups, and studied by means of Fluorine-19 NMR, as has myoglobin reconstituted with this porphyrin, allowing the study of various haem sites both in the isolated prosthetic group, and in a functionally unperturbed protein environment[126].

Xenon-129 NMR has been employed in studies on xenon binding to myoglobins and haemoglobins[127,128]. The Xenon-129 chemical shifts in a series of common solvents, oil, proteins and cell membranes have been compared, as has the rate of exchange of Xenon from the aqueous to the organic environments in such systems, in order to provide information on the anaesthetic action of noble gases[127]. The binding of Xenon by the cyanomet- and carbonmonoxy- forms of both myoglobin and

haemoglobin has also been monitored by Xenon-129 NMR[128].

Xenon adducts are of interest as Xenon occupies a unique hydrophobic cavity, the Xenon hole, within the protein matrix. Xenon is held in place in this site, which is located on the proximal side of the haem, abbutting the proximal histidine residue, purely by Van der Waals interactions with the surrounding groups. Furthermore Xenon appears to enter the protein molecule through a surface "gap" between the haem group, and leucine residues 86 and 89 and to approach the haem via a unique entry path requiring little or no rearrangement of the protein matrix. This is in direct contrast to the previously discussed entry channel employed by other ligands (eg. oxygen, carbon monoxide etc).

Phosphorus-31 chemical shifts of trimethyl phosphine bound to haemoproteins have been found to be very sensitive probes of the haem environment[129]. In ^{31}P NMR studies of trimethyl phosphine bound in various positions to sperm whale and equine myoglobin, and human, bovine, and rabbit haemoglobin, as well as various model compounds, the phosphorus-31 chemical shifts have been shown to be extremely sensitive to both the presence and the genetic variance of the globin, distinguishing, for example, between α and β chains of haemoglobins and seeming to bear a direct relationship to the facility of oxidation of the haemoprotein.

Various myoglobins and other haemoglobins substituted with prosthetic groups containing central metals other than iron have been studied. Proton NMR studies have been employed

to investigate the binding of the proximal histidine residue to the central porphyrin metal in cobaltomyoglobins and haemoglobins, and iron cobalt hybrid species[130,131], whilst Cadmium-113 NMR studies of Cadmium II substituted porphyrins and Cadmium II myoglobin have shown that the Cadmium-113 chemical shift, with a range of over 900ppm, affords a highly sensitive probe of both the coordination geometry of the central metal, and the nature of the coordinated ligand[132].

Recently Oxygen-17 NMR studies of carbonmonoxy haemoproteins have been reported[133], constituting the first observation of high resolution Oxygen-17 NMR spectra of the carbonmonoxy ligands of metalloproteins.

The Oxygen-17 spectra of enriched carbon monoxide ligated sperm whale myoglobin, and human and rabbit myoglobin have been recorded, and the results suggest a greater internal motional freedom of the ligand in haemoglobin than in myoglobin, which may have a bearing on the greater affinity of haemoglobin over myoglobin for carbon monoxide, reducing the magnitude of the sterically induced instability of the linear carbon monoxide molecule discussed earlier, in the case of haemoglobin, and thus increasing its carbon monoxide affinity with respect to myoglobin.

The study further shows that a direct linear relationship exists between the Oxygen-17 chemical shift of the bound carbon monoxide ligand and the carbon monoxide binding affinity of the protein under study.

CHAPTER 3

Experimental Procedures

"Finds tongues in trees,
books in running brooks,
Sermons in stones,
and good in everything"

William Shakespeare
As You Like It

3.1 Sample Preparation

Horse skeletal myoglobin (Type 1, 95-100%, crystalised and lyophilized) was purchased from Sigma, and used without further purification.

The paramagnetic low-spin met-cyano complex of myoglobin was prepared by dissolution of the lyophylate in D_2O (7mM), together with a trace of potassium ferri cyanide and a 50% molar excess of potassium cyanide. The pH value was adjusted to 8.4 by the use of NaOD and DCl, and a trace of dioxane was added to provide a secondary reference, designated 67.8 ppm for ^{13}C and 3.7 ppm for protons.

3.2 Instrumentation and Software.

3.2.1 Instrumentation and Software employed

All spectra were recorded on either Bruker AM300 or AM500 spectrometers, the AM300 operating at 300.13 MHz for protons and 75.47 MHz for carbon, and the AM500 at 500.14 MHz for protons and 125.75 MHz for carbon. Both spectrometers were equipped with ASPECT 3000 data systems, standard temperature control units, and Oxford Spectrospin cryomagnets, and ran Bruker DISB871 and DISR88 software respectively, and both employed Redfield quadrature detection[141]. Signals were detected using a Bruker 10mm multinuclear probehead, with deuterium lock and tunable X nucleus.

High quality 10mm Aldrich Gold label NMR tubes, suitable for high field studies, were used for all experiments.

3.2.2 Instrumentation Calibration

For optimum sensitivity the spectrometer probehead was tuned and matched for each experiment, using a Bruker reflection bridge. Experiments were performed in the temperature range 25°C-55°C, the experimental temperature being determined from the chemical shift of the methyl protons of isoleucine FG5.

3.3 Spectral Parameters

3.3.1 One Dimensional Spectra.

One dimensional spectra were recorded using a suitably large spectral width in order to avoid the folding back of outlying peaks, together with the collection of sufficient data points to give good digital resolution[140].

Typical proton spectra, for example, consisted of at least 400 transients collected on 16k data points of approximately 15KHz bandwidth at 300MHz.

Proton spectra were run with presaturation of the HOD signal in all cases.

3.3.2 Two Dimensional Spectra.

The digital resolution of the two-dimensional experiment is inherently poorer than that found in corresponding one

dimensional experiments due to time and data storage constraints, although this is compensated for by the increase in resolution which arises from the spread of signals into the second dimension.

For a typical two dimensional proton-proton shift correlation experiment 256 increments of t_1 were employed over 13.5KHz with 1024 data points collected in t_2, whilst a typical $^1H^{13}C$ shift correlation experiment might have a spectral width of 13KHz in the dimension of ^{13}C shifts, and an evolution period incremented in 256 steps of 152μs with a spectral width of 22ppm in the second (proton) dimension. The number of transients collected for each experiment was determined not only by the required signal to noise ratio, but also by the number needed to complete the phase cycling regime employed. A number of dummy scans were also inserted before the acquisition of the individual FID's in order to ensure that the system was at equilibrium.

Where necessary phase sensitive detection was employed in t_1, using time proportional phase incrementation (TPPI)[141] rather than the absolute value mode, (see section 3.3.3). This not only gives a better signal to noise ratio, but also retains valuable phase information on the individual cross peaks.

The apodization functions employed varied, but typically for homonuclear spectra sine bell functions were employed in both dimensions, whilst simple exponential window functions were utilized in the case of heteronuclear spectra.

Undesired artefacts in homonuclear two dimensional

spectra were reduced by means of the spectrometers symmetrization routine, as signals which are non-symmetrical about the diagonal are likely to be artefacts. This procedure, however, is not applicable to the heteronuclear two dimensional spectra, which possess non diagonal symmetry and are consequentially more prone to spurious noise.

Two dimensional spectra are presented as contour level diagrams. The diagonal in homonuclear spectra may be considered as an approximation of the one dimensional spectrum, with cross peaks representing J-coupled partners, whilst the individual peaks represent discrete structural units (CH_n) in the heteronuclear spectra. Each contour line represents a factor of $\sqrt{2}$ difference in intensity from the next level unless otherwise stated.

3.3.3 Phase Sensitive Detection and Phase Cycling

In order that more effective radiation over the range of frequencies involved may be effected, it is desirable that the carrier frequency ν_o be placed at the centre of the spectrum, however, this requires a detector set up which can distinguish not only a difference in magnitude between the oscillator and the detected signal, but can also distinguish the sign of that difference in frequency. This is achieved by quadrature detection whereby two detectors are employed, one of which is 90° out of phase with the oscillator. Consequently the sine and cosine components of the magnetisation may be simultaneously

detected. A complex fourier transform then gives the relative signs of the signals with respect to the central carrier frequency and consequently only frequencies up to ±F/2 need to be characterised[140].

The method of quadrature detection used, in practice, by Bruker spectrometers is the Redfield method[141] which can be regarded as a pseudo quadrature detection method.

Quadrature detection can unfortunately give rise to artefacts, known as quadrature images in the spectrum. These take the form of small images of the main peak symmetrical about the centre of the spectrum, and are due to imperfect definition of the relative sign of the frequencies.

These images may be averaged out by means of the cycling of the relative phase of the transmitter, with respect to the detector, over the different axes such that the unwanted components of the magnetisation are cancelled.

3.4 Analysis of Spectral Data.

An outline of the NMR methodologies employed in this work is presented, together with the kind of information which may be derived there from. The relative advantages and disadvantages of certain of the techniques are also explored.

3.4.1 One Dimensional Spectra.

Inspection of one dimensional proton and carbon spectra (Figs 4.iv and 4.v) shows that distinct spectral regions may be recognised. Indeed, as outlined previously, much valuable information may be gleaned from chemical shift data alone[21,26], particularly in the case of those resonances of paramagnetic species, such as met-cyanomyoglobin, which experience significant hyperfine shifts due to their proximity to the paramagnetic centre. Furthermore, the temperature dependence of these paramagnetically shifted resonances may be employed as both an aid to spectral resolution and as a source of additional data (see section 4.2.1).

The data obtainable from a one dimensional ^{13}C spectrum may be further refined by editing of the spectrum according to the number of attached protons. Carbon-13 resonances arising from methyl, methylene and methine groups may be separated, by magnetisation transfer from their attached protons, by means of the DEPT pulse sequence[144]. A final, variable length, proton pulse, θ, in the sequence leads to different mathematical dependencies for different CH_n groups, such that the intensity of the CH varies with sin θ, CH_2 with sin 2θ and CH_3 with sin θ + sin 3θ (see fig 3.i), consequently individual CH, CH_2 and CH_3 sub spectra may be generated by means of appropriate combinations of spectra recorded with different values of θ.

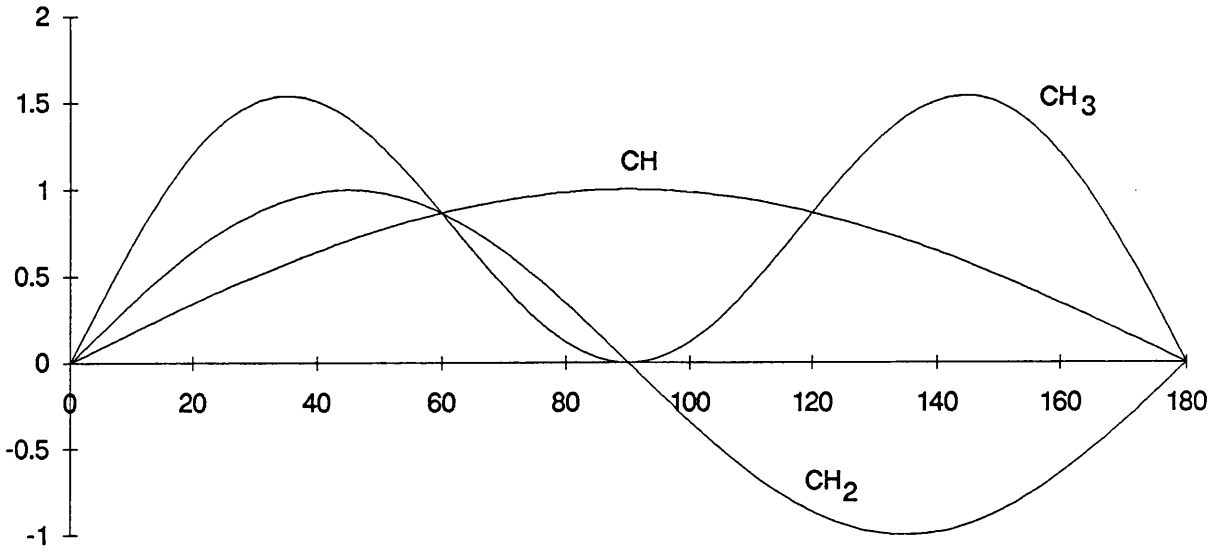


Fig. 3.i

The variation in intensity of CH_n groups with θ pulse length.

3.4.2. Two Dimensional Spectra.

The homonuclear shift correlation experiment, COSY[33], is mediated by scalar J coupling which is a through bond effect, transmitted between nuclei via the electrons of intervening bonds, and occurs through the pairing of electronic and nuclear spins. The magnitude of this coupling, which is a field independent parameter, is related to the type and number of intervening bonds, and falls away rapidly with increasing bond

number. The cross peaks represent J coupled partners within the spectrum, allowing the characteristic intra residue J coupling patterns of individual amino acid residues to be mapped out. An example of a COSY spectrum is shown in figure 4.ix.

This two-dimensional methodology for the identification of J coupled partners has two main advantages over the one dimensional alternative of homonuclear double-resonance experiments, namely that all partners in a coupling network may be observed simultaneously, and that potentially crowded spectral regions in the first dimension are spread out in the second dimension. This second advantage also permits the tracing of the temperature dependencies of resonances which are unresolved in one dimensional spectra.

The standard pulse sequence of the basic COSY experiment is given in scheme 3.i.

$$90^\circ \quad -t_1- \quad 90^\circ \quad -FID(t_2)$$

Scheme 3.i.

Standard Pulse Sequence of the COSY Experiment

Simple modifications of this pulse sequence allows optimisation for longer range coupling constants.

The insertion of a fixed time delay, τ, at the beginning of the t_1 and t_2 periods of the sequence emphasizes long range coupling constants by allowing these magnetisation components, whose intensities increase with $\sin(\pi J\tau)$, more time to evolve

for maximum sensitivity, whilst simultaneously suppressing the diagonal peaks[145]. This allows identification of longer range coupling than would the standard COSY sequence.

Similarly COSY-45 employs a 45° mixing pulse, rather than the standard 90° pulse, which results in a "streamlining" of the diagonal, clarifying some of the more complex regions of the spectrum[145].

Heteronuclear shift correlation experiments allow the consideration of individual CH_n groups as discrete spectroscopic units, and combine the advantages of two dimensional spectroscopy with that of the greater chemical shift dispersion of the carbon spectrum with respect to the proton spectrum.

The technique employed for the correlation of proton and carbon-13 shifts utilizes polarisation transfer via scalar J couplings. This gives rise to an additional advantage in terms of the carbon-13 dimension, in that a polarisation transfer from ^1H to ^{13}C takes place, with a resultant four fold increase in carbon-13 sensitivity. The pulse sequence is shown in scheme 3.ii.

^1H	90°-$t_{½}$-		-$t_{½}$-τ_1-90°-	BB
^{13}C		180°	90°-τ_2	FID(t_2)

Scheme 3.ii Pulse Sequence for $^1H^{13}C$ Shift Correlation

It can be seen from examination of scheme 3.ii that this sequence is related both to the COSY experiment[33], and to the INEPT experiment described by Morris and Freeman[146].

As is the case with the parent INEPT experiment, the intensity of a signal in the heteronuclear shift correlation experiment varies not only with the length of the final delay, τ_2, but also with its multiplicity, and $^{13}C-^{1}H$ coupling constant.

In the normal case the τ_2 period is selected as a compromise for the various multiplicities and J values, at a value of around 1/4J, see for example figure 4.vi, although optimum τ_2 selection is hampered by the variability of J values.

It is also possible however, by appropriate choice of delay length, to generate $^{1}H^{13}C$ shift correlation spectra which are selective for signal multiplicity, thereby conferring an additional advantage on the experiment. Thus by choosing τ_2 = 1/2J a spectrum selective for methine groups may be obtained, whilst Fig. 4.xv shows a spectrum in which signals due to methyl and methine groups give rise to positive cross peaks, and methylene groups gives rise to negative cross peaks.

CHAPTER 4

Results and Discussion

"If the doors of perception were cleansed,
everything would appear to man as it is,
... infinite."

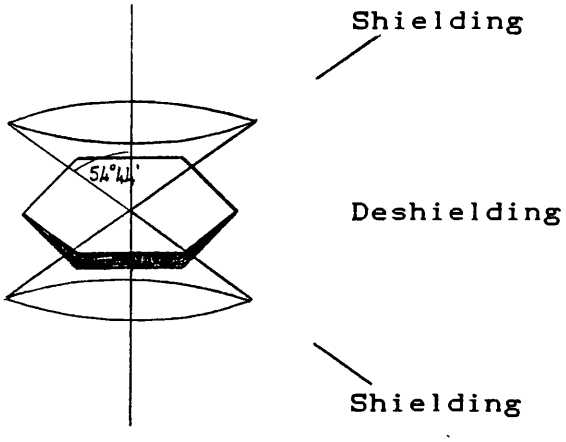
William Blake
The Marriage of Heaven and Hell

4.1 Particular Factors Influencing the NMR Spectra of Proteins.

Certain factors, such as the "ring currents" present in aromatic systems, and the hyperfine interactions due to the presence of a paramagnetic ion, which influence the chemical shifts of individual resonances, are of particular importance in the NMR spectra of haemoproteins, and these are discussed below.

4.1.1 Ring Current Effects

The circulation of delocalised π-electrons in the plane of an aromatic ring under the influence of an external magnetic field is known as a ring current, and produces a further local magnetic field. This opposes the externally applied field in the areas above and below the plane of the ring, but reinforces it in all other areas. (Fig 4.i).



"Ring current effects in a benzene ring" (Fig 4.i)

Quantitative calculations of the magnitude of the shielding zone about the benzene ring have been performed by Johnson and Bovey[155]. In the case of proteins ring currents are associated with the delocalised electrons arising from the aromatic amino acid residues, histidine, phenylalanine, tryptophan and tyrosine. Additionally ring currents in proteins may arise due to the presence of aromatic prosthetic groups such as porphyrins.

The ring current fields associated with the porphyrin rings[156] are much larger than those associated with simple aromatic systems. It is estimated that the resonance of a proton near the plane of a phenylalanine ring for example may experience an upfield shift of up to around 2ppm depending on its exact position, whilst a proton near the plane of a haem group can experience an upfield shift of in excess of 5ppm[75].

4.1.2 Hyperfine Shifts.

The presence of a paramagnetic centre can also effect the chemical shifts of resonances. The magnetic resonances of nuclei in paramagnetic complexes often show large shifts from their corresponding positions in diamagnetic complexes. These shifts are known as hyperfine shifts and are often accompanied by paramagnetic broadening of the resonances. Hyperfine shifts, which may be as large as several hundred parts per million, are caused by interactions with the unpaired electrons of paramagnetic species, and arise in two different ways, namely contact and pseudocontact or dipolar shifts.

Contact shifts sometimes called "Fermi" or isotropic

contact interactions, result from the transfer of finite amounts of electron spin density from the unpaired electron(s) of the paramagnetic species to the observed protons. This normally occurs as a through bond transfer, and the contact shifts of proton resonances for isotropic systems are given by Bloembergens equation (Eqn 4.1), where A_i is the contact

$$\Delta\nu_{ci} = -A_i \frac{\gamma_e}{\gamma_H} S(S+1) \frac{\nu}{3kT} \qquad \text{.....Eqn 4.1}$$

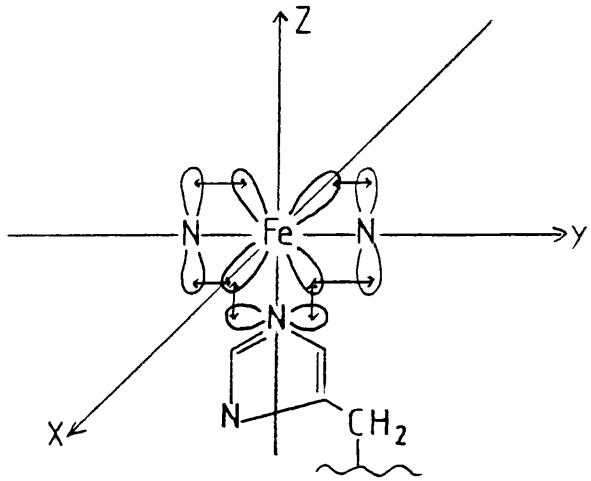
interaction for the ith proton, γ_e and γ_H are the magnetogyric ratios of the electron and proton respectively, S is the total electronic spin, ν is the resonance frequency of the proton, k is the Boltzmann constant, and T is the absolute temperature.

In the haems and haemoproteins contact shifts arise when finite amounts of electron density are delocalised from the iron orbitals into the π-orbital system of the porphyrin and the axial ligands as indicated in figure 4.ii. The electron density is then further transferred from the aromatic ring carbons to the protons, giving rise to contact interactions.

Pseudocontact shifts, sometimes known as dipolar shifts, arise when the electronic g-tensor is anisotropic, that is when the dipolar interaction between the unpaired electron spin and the nucleus fails to average zero.

Pseudocontact shifts are related to the value of the

electronic g-tensor, which is a measure of the field experienced by the unpaired electrons.



(Fig 4.ii)

Interactions of the $3d_{yz}$ ($3d_{xz}$) atomic orbital of the haem iron with the π-orbitals of the ligand nitrogen atoms[75].

If certain assumptions are made then the pseudocontact shifts may be obtained from a knowledge of the molecular geometry and of the principal values of the electronic g-tensor, which may be derived from electron paramagnetic resonance studies, by use of the equation 4.2 [75]:-

$$\Delta_{vpci} = \frac{-\beta^2 \nu S(S+1)(3\cos^2 X_i - 1)}{3kT} \frac{1}{r_i^3} \frac{(g_{\parallel} + g_{\perp})(g_{\parallel} - g_{\perp})}{3} \quad \text{...Eqn. 4.2}$$

Where β is the Bohr magneton, r_i is the distance from the haem iron to the observed proton, X_i is the angle between the vector r_i and the g axis and $g_{\parallel}$ and $g_{\perp}$ are the parallel and the perpendicular components, respectively, of the electronic g-tensor. In a dilute solution of a paramagnetic species the

transverse nuclear relaxation may be totally dominated by coupling between nuclear and electronic spins, both by dipole dipole mechanisms and by scalar coupling transmitted via, possibly transient bonds and line broadening will thus result, such that, in general, proton NMR spectra are likely to yield information on paramagnetic molecules only if the electron spin relaxation time is short, such that the hyperfine shifts are large compared with the line widths.

4.2 Assignments of the Haem Resonances.

4.2.1 Haem Methyl Groups.

The four haem methyl groups, borne on the porphyrin ring of haemoproteins have long been the subject of NMR studies, being uniquely sensitive to perturbations of the haem electronic structure.

The magnitudes of their individual hyperfine shifts have been shown to depend primarily upon spin density in the partially filled π orbitals of the pyrrole ring system[147] and studies of model systems in which hemins are chelated with various imidazoles of increasingly fixed geometry[148] have shown that the increased asymmetry of the haem methyl hyperfine shifts observed in low spin ferric cyanide complexes of haemoproteins over that found in the free protoporphyrin compounds is due to the interaction of the $p\pi$ orbitals of the proximal imidazole system and the metal d orbitals (π d_{xz},

d_{xy}). This in turn leads to a redistribution of spin density within the porphyrin system, such that the unpaired spin density is at right angles to the imidazole plane[148].

The orientation of the imidazole plane in myoglobin is shown in Figure 4.iii.

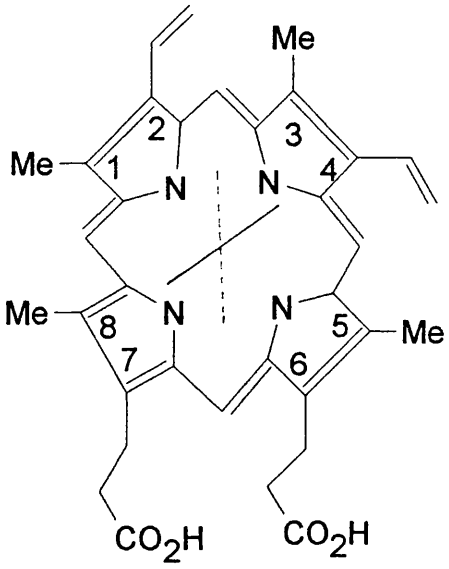


Figure 4.iii

Orientation of the Imidazole Plane in Myoglobin (———) and Cytochrome c (-----) as described in Reference 148.

Overlap with this orbital should result in pyrrole substituents located on an axis orthogonal to this plane experiencing the greatest hyperfine shifts, thus correctly predicting the large shifts displayed by the 1-methyl and 5-methyl groups in metmyoglobin cyanide. The inversion of this

pattern of haem methyl hyperfine shifts has been used[6] to solve the structure of the minor isomer of myoglobin, in which the haem group is rotated by 180° about the α-γ meso axis. The NMR behaviour of the pyrrole substituents can consequently be used as a probe for haem-imidazole geometry both in known, and in less well characterised haemoproteins, so it is therefore of supreme importance that a full and unambiguous set of assignments exist for all four haem methyl groups.

The proton resonances of the haem methyl were amongst the first to be recognised in protein NMR spectroscopy, and assignments were published in 1974 for those three methyl groups of sperm whale myoglobin; the 5-, 1-, and 8- methyl groups, whose resonances where shifted downfield of the diamagnetic envelope[74], and cross assignments into the carbon-13 spectrum have subsequently been made by means of $^{1}H^{13}C$ shift correlation spectroscopy. (see Table 2.ii.).

The hyperfine shift experienced by the final haem 3-methyl group, however, is insufficient to distance it from the diamagnetic envelope of the proton spectrum, (see Figure 4.iv- Proton NMR spectrum of Equine Met-Myoglobin cyanide.) and its unambiguous assignment consequently remained unsolved when this work was carried out, although two contradictory assignments had been published, the first assigning the proton resonance at 5.1ppm[108], and the second assigning the carbon-13 resonance at -28.5ppm (see Figure 4.v - Carbon-13 NMR Spectrum of Equine Met-Myoglobin cyanide). and consequently cross assigning the proton resonance, by two-dimensional shift correlation

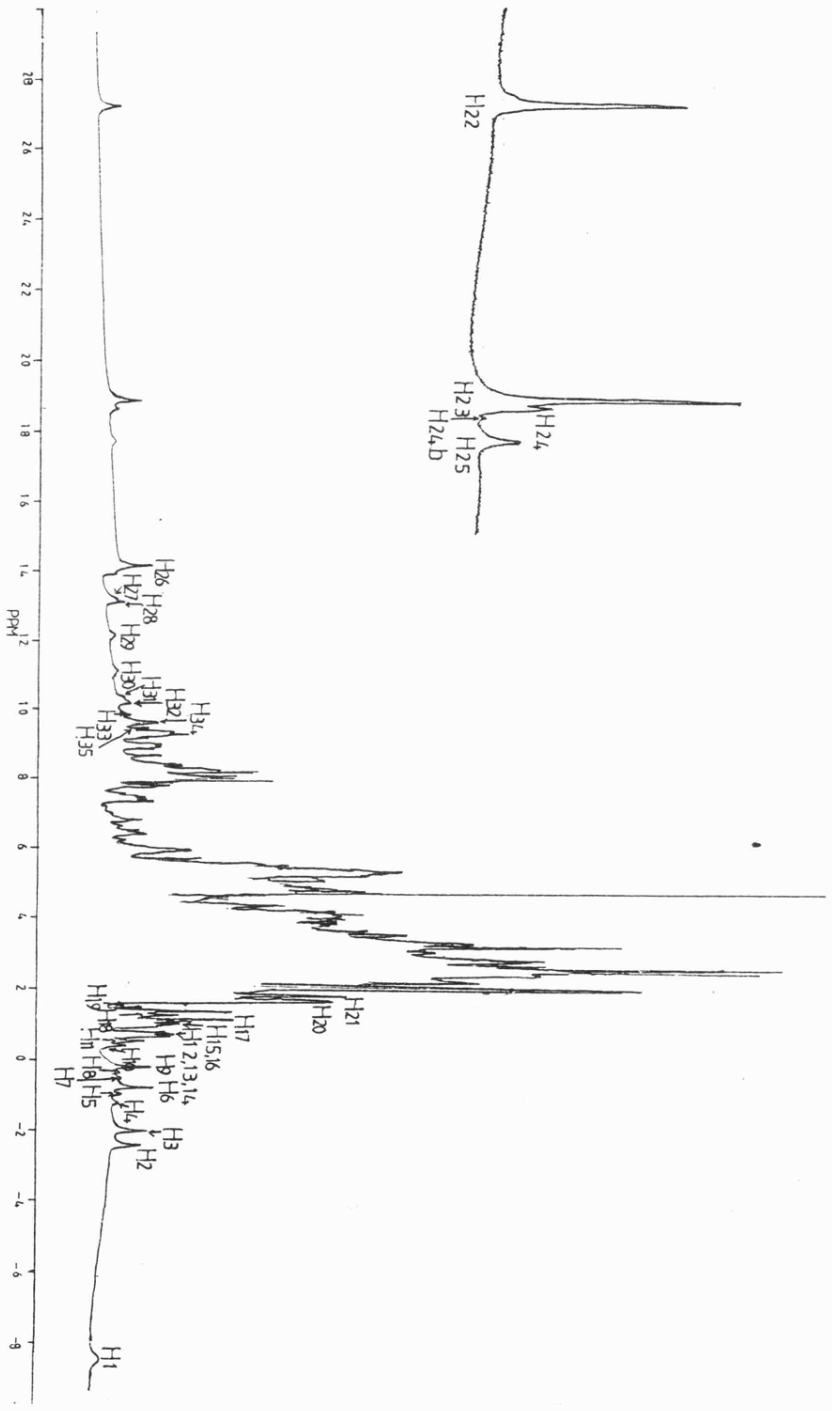


Figure 4.iv.- Proton NMR spectrum of Equine Met-Myoglobin cyanide in D_2O, 300MHz, 310K.

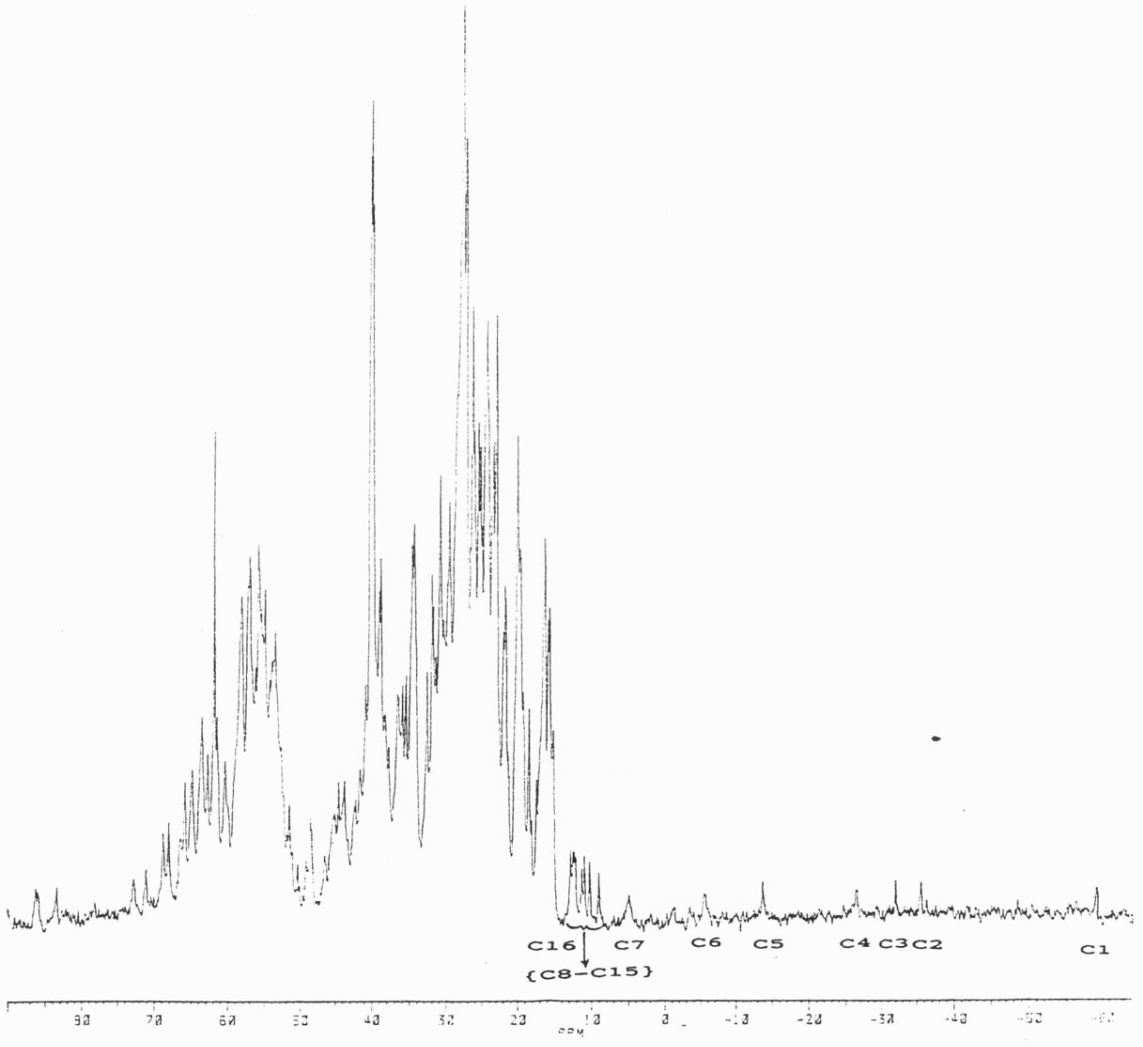


Figure 4.v. -Section of the Carbon-13 NMR Spectrum of Equine Met-Myoglobin cyanide in D_2O, 300MHz, 310K.

spectroscopy, at -1.7ppm[134].

Studies were therefore undertaken to determine which, if either, of these two conflicting results was the case.

Comparison of the proton and carbon-13 spectra of equine metmyoglobin cyanide (Figures 4.iv. and 4.v.) with the two dimensional proton-carbon-13 shift correlation spectrum (Figure 4.vi.) clearly shows the presence of the three signals at the appropriate chemical shifts for the haem 5-, 1-, and 8-methyl groups (labelled M1, M2 and M3 respectively). No cross peak is observed to be manifested, however, by the carbon signal, C-4, located at -28ppm, whereas a haem methyl group would be expected to give rise to as strong cross peak.

By editing the natural abundance carbon-13 spectrum according to the number of attached protons, using the method of Doddrell et al[144], individual sub-spectra of methyl, methylene and methine groups may be generated.

Reference to a carbon-13 spectrum of equine met-myoglobin cyanide thus edited (Figure 4.vii.) clearly shows that the carbon signal at -28ppm does not arise from a methyl group. A strong methyl resonance is observed, however, at ~ -14ppm, equivalent to signal C-5 in the carbon-13 spectrum (Figure 4.v.).This signal exhibits a strong cross peak, labelled M4, in the proton-carbon-13 shift correlation spectrum, with its corresponding proton chemical shift located at ~ 5.1ppm.

Furthermore studies of equine heart ferricytochrome c[59a] reveal a pattern of cross peaks in the $^{1}H^{13}C$ shift correlation spectrum very similar to that observed here for equine

cyanometmyoglobin, in which the cross peak equivalent to cross peak M4 arises from a haem methyl group.

Examination of equations 4.1 and 4.2 illustrate that both Δ_{vc} and Δ_{vpc} are proportional to the reciprocal of the absolute temperature. Ring current shifts, however are intrinsically temperature independent, provided that no temperature related changes in molecular conformation occur[157]. A plot of chemical shifts against the reciprocal of the absolute temperature [1/T(K)], known as a Curie plot, should therefore give a series of straight lines, displaying zero gradient, in the case of ring current shifted resonances, but showing some temperature dependence (gradient) in the case of any hyperfine shifted resonances.

Equations 4.1 and 4.2 also imply that as 1/T(K) approaches zero the hyperfine shifts should also approach zero. Thus the plots of chemical shift versus 1/T(K) should tend towards the chemical shifts of the corresponding resonances in the related diamagnetic compounds, when extrapolated to infinite temperature (ie 1/T(K) = 0). It follows that where a difference in the relative shifts of a group of resonances is due almost exclusively to differing hyperfine interactions, as, for example, in the case of the four ring methyls of a haemoprotein, the plots should converge towardsa common diamagnetic shift at infinite temperature. In practice however, whilst these phenomena occur, their incidence appears to be at somewhat lower temperature than that expected and, in particular in the case of signals which experience large

pseudocontact shifts, may have extrapolated intercepts which tend to overshoot their expected diamagnetic chemical shifts.

Wüthrich[75] has suggested that such observed deviations from Curie behaviour are due to the fact that in the temperature range of the NMR experiment the scalar coupling constant (A in equation 4.1) is itself temperature dependent, whilst La Mar et al [153] explain the effect as being due to the contribution to the overall equation of the paramagnetic shift terms, which are quadratic with respect to inverse temperature. Large deviations from the expected variable temperature behaviour of paramagnetic shifts may also occur in species in which multiple spin states exist in equilibrium[154]

If the signal M4 is indeed due to the remaining haem methyl group therefore it should be expected to exhibit a strong temperature dependency in its chemical shift, and furthermore, a plot of chemical shift against inverse temperature (Curie Plot) should give a straight line which extrapolates to the diamagnetic shift at infinite temperature.

Examination of this variable temperature behaviour (Figure 4.viii.) shows that the four methyl signals do indeed converge towards a diamagnetic intercept.

I therefore assign cross peak M4 to the remaining haem-3-methyl group. This completes the assignment of the haem methyl groups in both the proton and carbon-13 spectra. The full assignments are presented in Table 4.i.

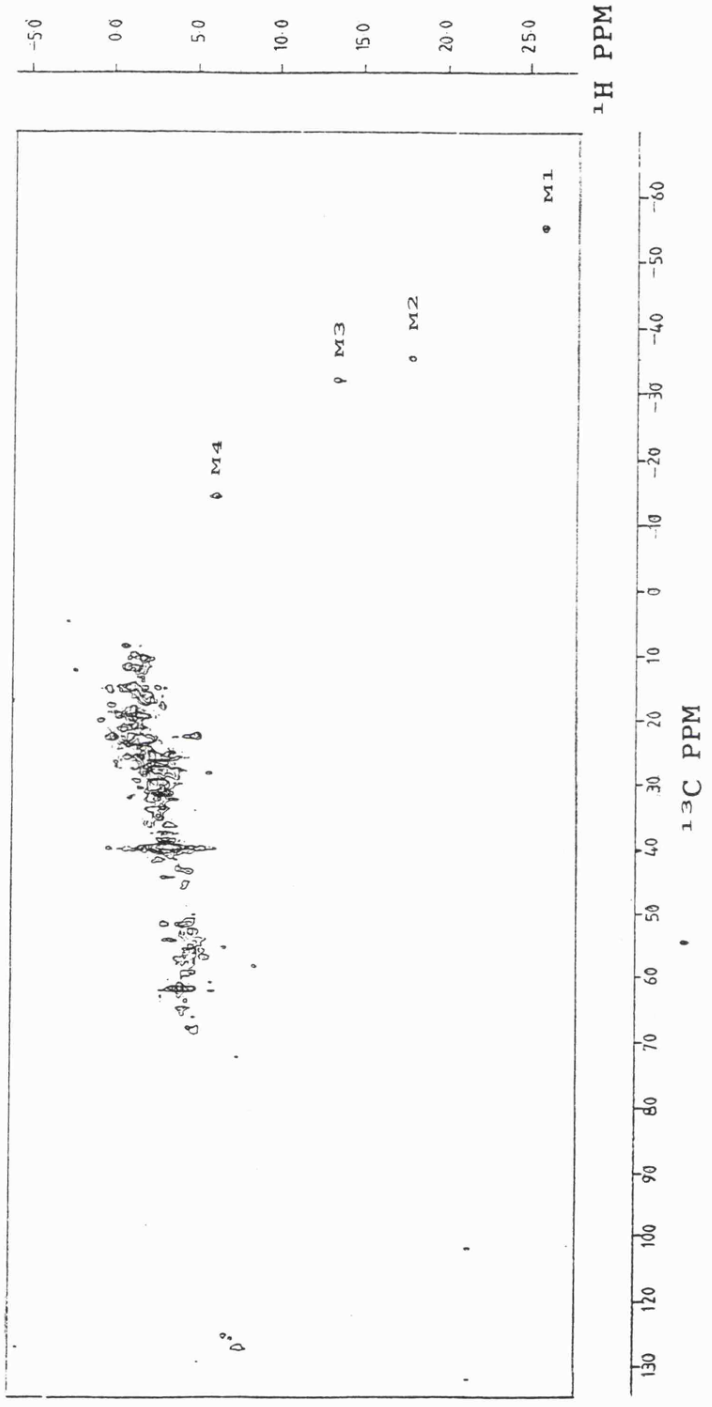


Figure 4.vi $^{1}H^{13}C$ Shift Correlation Spectrum of Equine met-myoglobin cyanide in D_2O, 300MHz, 303K.

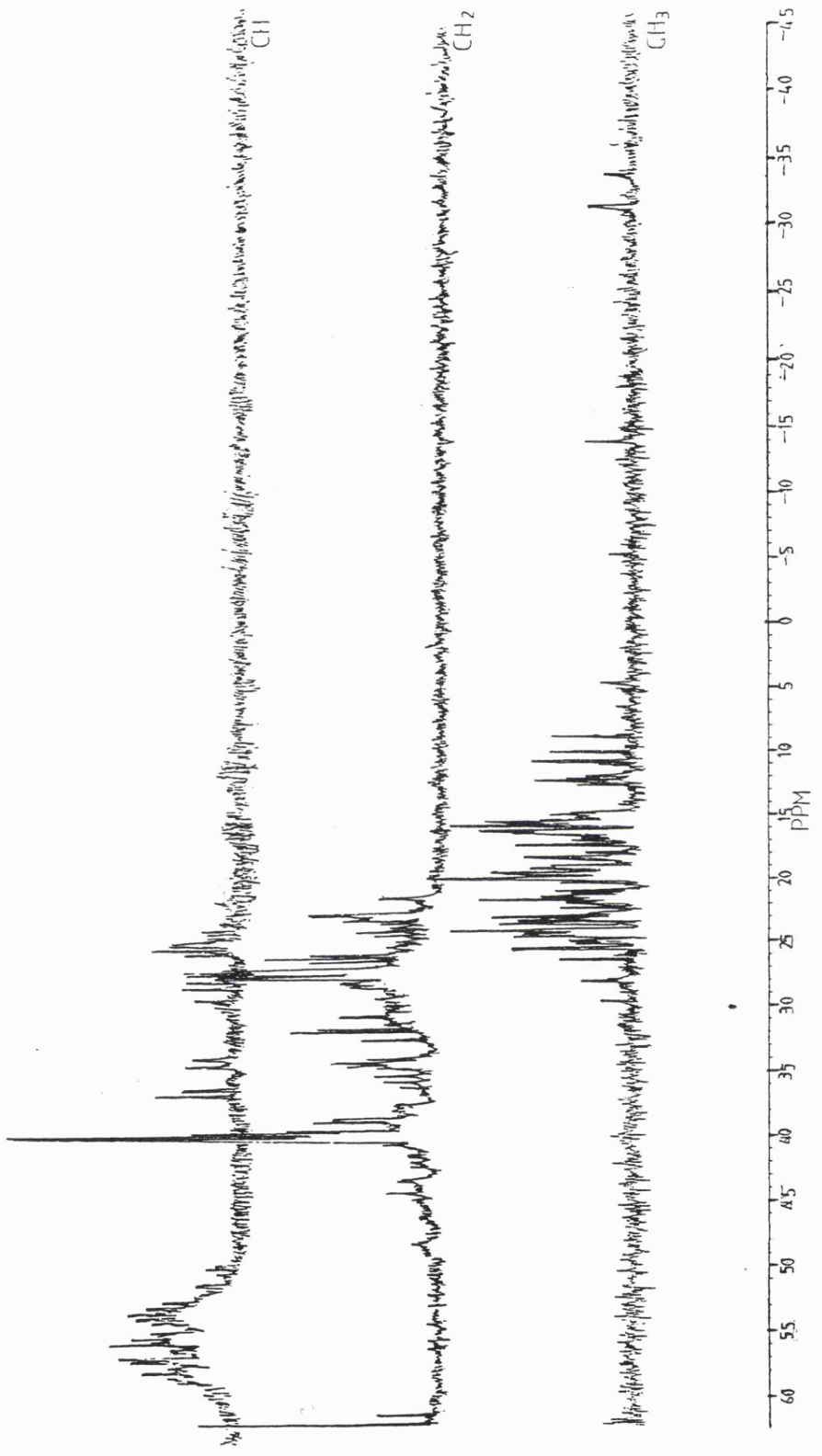


Figure 4.vii Part of the ^{13}C NMR Spectrum of the Equine Met-myoglobin cyanide 310K, separated into signals arising from methyl, methylene and methine groups.

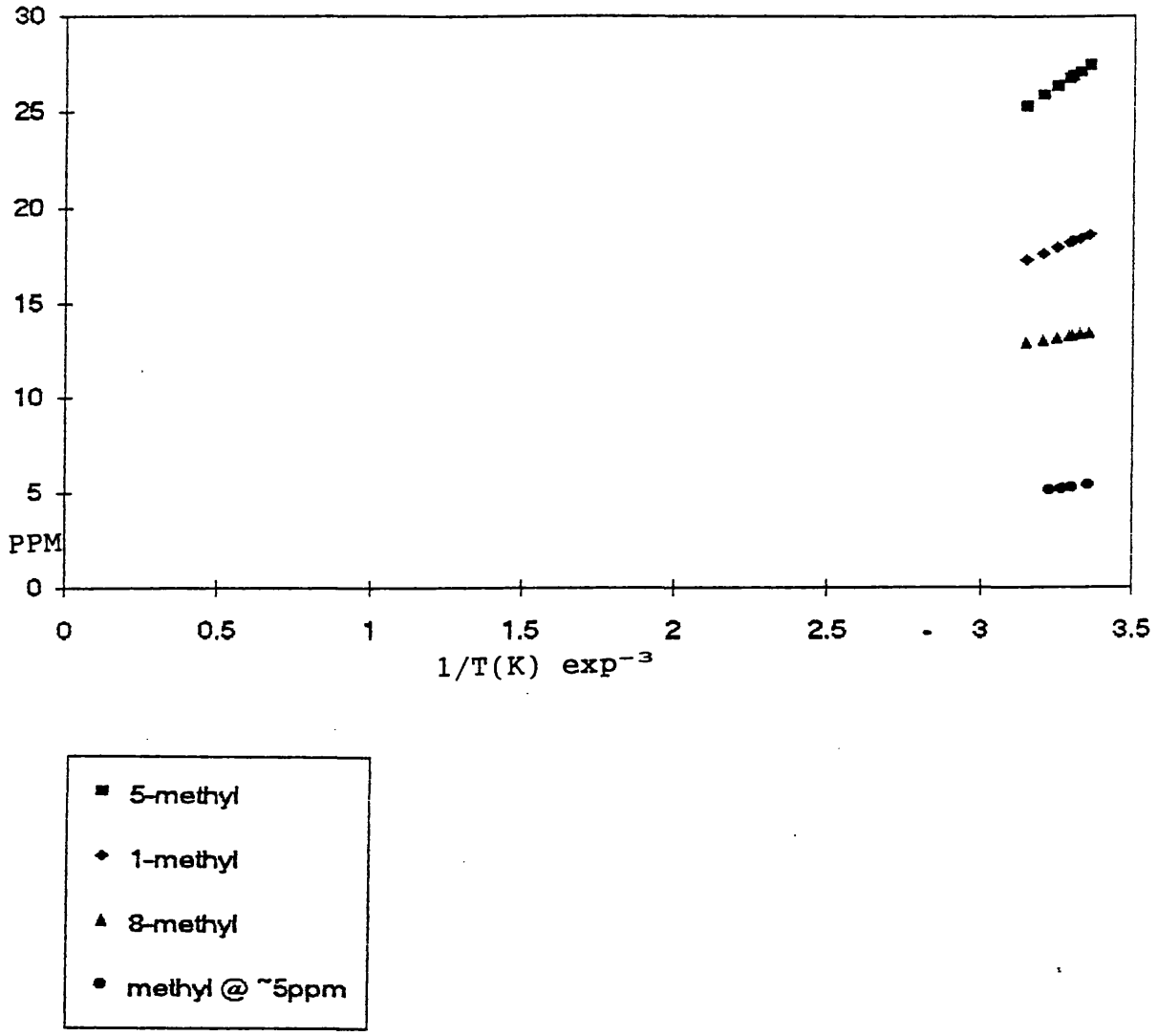


Figure 4.viii "Proton Chemical Shifts of the three known and the fourth proposed, haem methyl groups plotted against inverse temperature".

Haem Methyl Group	^{1}H Chemical Shift	^{13}C Chemical Shift
5	26.07[b]	-59.01[c]
1	17.54[b]	-35.43[c]
8	13.06[b]	-31.94[c]
3	5.09	-13.64

Table 4.i.

Chemical Shifts of Haem Methyls of MetMbCN

a. Shifts in ppm from TMS, D_2O solution @ 310K

b. Reference 74.

c. Reference 134.

The spread of the haem methyl proton resonances reflects the asymmetric nature of the haem electronic structure, and therefore the spread of the literature values for sperm whale cyanometmyoglobin,[74,108] was compared with that obtained herein for equine cyanometmyoglobin, at the same temperature :-

Sperm Whale 3Me 5.1ppm, 8Me 13.0ppm, 1Me 18.7ppm, 5Me 27.4ppm.
Mean 16.05ppm, Spread 22.3ppm.

Equine 3Me 5.09ppm,8Me 13.43ppm,1Me 18.56ppm,5Me 27.4ppm.
Mean 16.12ppm, Spread 22.31ppm.

The close correspondance of these values suggests a near identical haem electronic environment in these two genetic variants.

Spectra of model compounds[61a,75,134] show that the α carbons and attached protons of the haem 6- and 7- propionate groups also experience significant upfield shifts, bringing them into the same spectral region as the haem methyl signals, and these would be unlikely to give rise to strong cross peaks in the $^{1}H^{13}C$ shift correlation spectrum due to the inequivalence of the diastereotopic propionate α methylene proton resonances. Furthermore the attached 6 and 7α propionate protons have been reported[85] for sperm whale myoglobin as :-

6Hα ~9.2ppm; 6Hα' ~7.4ppm; 7Hα ~1.1ppm; 7Hα' ~-0.5ppm

and I therefore tentatively assign the resonance at -28ppm (^{13}C), -1.7ppm (^{1}H) as the haem 7CHα'.

4.2.2 Haem Vinyl Groups.

The highly shifted resonances of the haem 2-vinyl group, which are fully resolved from the diamagnetic envelope in the one dimensional proton spectrum, have been identified and assigned[87] (see Table 2.ii) and are designated in Figure 4.iv. as :-

$H24 = 2C_{\alpha}H$ $H4 = 2C_{\beta}H_{trans}$ $H8 = 2C_{\beta}H_{cis}$

Table 4.ii

Chemical Shifts of Resolved Proton Resonances in the 1D Spectrum of Equine Cyanometmyoglobin at Various Temperatures.

Temp. (°C)	25	28	30	31	35	39	45
Resonance							
H1	-9.489	-9.213	-9.090	-9.009	-8.740	-8.377	-7.912
H2	-3.758	-3.664	-3.606	-3.584	-3.468	-3.337	-3.163
H3	-3.322	-3.228	-3.177	-3.163	-3.061	-2.937	-2.785
H4	-2.589	-2.480	-2.422	-2.407	-2.277	-2.102	-1.942
H5	-2.489	-2.298	-2.218	-2.182	-2.030	-1.833	-1.623
H6	-2.030	-1.964	-1.921	-1.913	-1.833	-1.732	-1.623
H7	-1.928	-1.812	-1.754	-1.717	-1.601		-1.456
H8	-1.616	-1.543	-1.500	-1.470	-1.405	-1.282	-1.165
H9	-1.412	-1.354	-1.325	-1.311	-1.238	-1.158	-1.056
H10	-0.679	-0.657	-0.650	-0.650	-0.621	-0.599	-0.548
H11	-0.497	-0.483	-0.483	-0.483	-0.475	-0.483	-0.468
H12	-0.395	-0.381	-0.366	-0.361	-0.352	-0.352	-0.316
H13	-0.308	-0.316	-0.308	-0.316	-0.308	-0.301	-0.279
H14		-0.272	-0.265	-0.265	-0.250	-0.228	-0.199
H15	(-0.003)	-0.156	-0.127	-0.120	-0.069	(-0.011)	-0.011
H16	-0.003	-0.011	-0.011	-0.003	-0.004	-0.011	-0.004
H17	0.069	0.091	0.099	0.099	0.113	0.120	0.157
H18	0.287	0.280	0.273	0.266	0.266	0.251	0.251
H19	0.338	0.353	0.345	0.338	0.345	0.360	0.367
H20	0.600	0.614	0.614	0.614	0.621	0.636	0.650
H21	0.767	0.774	0.774	0.767	0.781	0.759	0.774
H22	27.400	27.095	26.877	26.746	26.383	25.860	25.293
H23	18.561	18.372	18.256	18.154	17.934	17.609	17.275
H24	17.900	17.834	17.791	17.733	17.667	17.544	17.420
H25	17.500	17.304	17.181	17.050	16.825	16.432	16.040
H26	13.433	13.368	13.317	13.281	13.186	13.041	12.903
H27	12.554	12.503	12.402	12.351	12.235	12.046	11.873
H28	12.416	12.344	12.278	12.235	12.140	11.843	11.305
H29	11.654	11.530	11.457	11.363	11.232	10.971	10.746
H30	10.731	10.586	10.477	10.375	10.201	9.932	9.656
H31	(9.279)	(9.271)	(9.257)	(9.235)	9.279	9.279	9.293
H32	9.279	9.271	9.257	9.235	9.213	9.148	9.126
H33	8.952	8.937	8.930	8.894	8.870	8.821	8.785
H34	8.661		8.690	8.669	8.669	8.625	8.611
H35	8.729	(8.937)	8.690	8.669	8.631	8.574	8.538

Note. Figures in brackets indicate peaks unresolved from their immediate neighbours at that temperature.

and the expected COSY cross peaks, $2\alpha H\text{-}2\beta H_{cis}$, and $2\alpha H\text{-}2\beta H_{trans}$ are identified as x and y respectively in the two dimensional

proton-proton COSY spectrum of cyanomet-myoglobin. (Figure 4.ix).

The 4-vinyl groups, however, being situated closer to the plane of the proximal imidazole than the 2-vinyl group, experiences a lesser hyperfine shift and its resonances are unresolved in the one dimensional proton spectrum and remain unassigned.

Examination of the proton-proton COSY spectrum (Figure 4.ix) reveals a pair of cross peaks, similar to those observed for the 2-vinyl group designated l and m, linking a resonance at 5.8ppm with resonances at -0.57ppm and -1.8ppm, this latter resonance being equivalent to peak H7 in the one dimensional proton (Figure 4.iv). These seem the most likely candidates for the haem vinyl resonances, due to their characteristic coupling pattern, and to the spectral window which they occupy. Furthermore the chemical shifts of these resonances display a marked dependence on experimental temperature, as would be expected for substituents of the porphyrin ring. (see tables 4.ii, and 4.iii, and Figure 4.x).

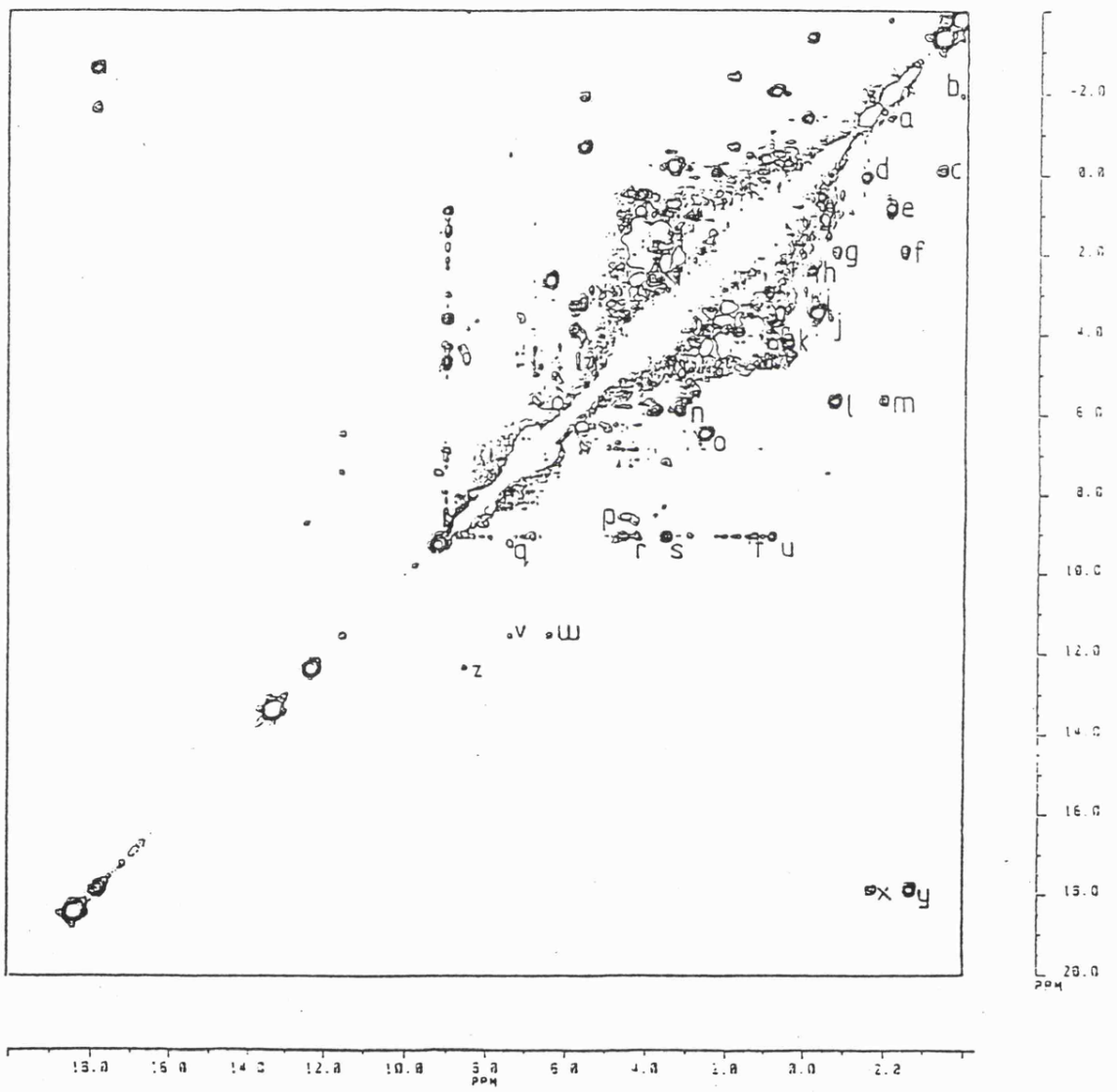


Figure 4.ix

Two dimensional proton-proton COSY spectrum of the equine skeletal cyanomet-myoglobin.

Table 4.iii

Variation with Temperature of the Chemical Shifts, in ppm of Cross Peaks l,m,x and y in the $^1H^1H$ COSY Spectrum of Equine Skeletal cyanomet-myoglobin.

Temp.(°C)	27	29	30	33	35
Peaks	5.67	5.67	5.76	5.76	5.84
l and m	-1.80	-1.80	-1.75	-1.71	-1.54
	-0.57	-0.57	-0.57	-0.57	-0.44
Peaks	17.90	17.81	17.72	17.72	17.64
x and y	-2.51	-2.51	-2.42	-2.42	-2.24
	-1.54	-1.54	-1.54	-1.54	-1.45

Temp.(°C)	37	39	40	42	47	52
Peaks	5.85	5.85	5.94	6.02	6.29	6.38
l and m	-1.54	-1.50	-1.28	-1.28	-1.10	-0.83
	-0.44	-0.40	-0.40	-0.13	0.04	0.13
Peaks	17.72	17.63	17.63	17.46	17.02	17.10
x and y	-2.22	-2.22	-2.07	-1.98	-1.80	-1.63
	-1.28	-1.28	-1.36	-1.36	-1.01	-0.83

Curie plots of these resonances would be expected to give rise to intercepts at their chemical shifts in the corresponding diamagnetic species, and the proton chemical shifts of the haem vinyl resonances are given by Dalvit and Wright[149] as :-

$2H\alpha$	~ 8.43ppm	$4H\alpha$	~ 8.62ppm
$2H\beta_{cis}$	~ 5.69ppm	$4H\beta_{cis}$	~ 6.29ppm
$2H\beta_{trans}$	~ 5.73ppm	$4H\beta_{trans}$	~ 6.61ppm

The Curie plots for these six resonances are given in figure 4.xi., and by comparison of the known intercepts I assign the four vinyl resonances as follows

$4H\alpha$ = 5.8ppm; $4H\beta_{trans}$ = -0.6ppm; $4H\beta_{cis}$ = -1.8ppm (@ 303K)

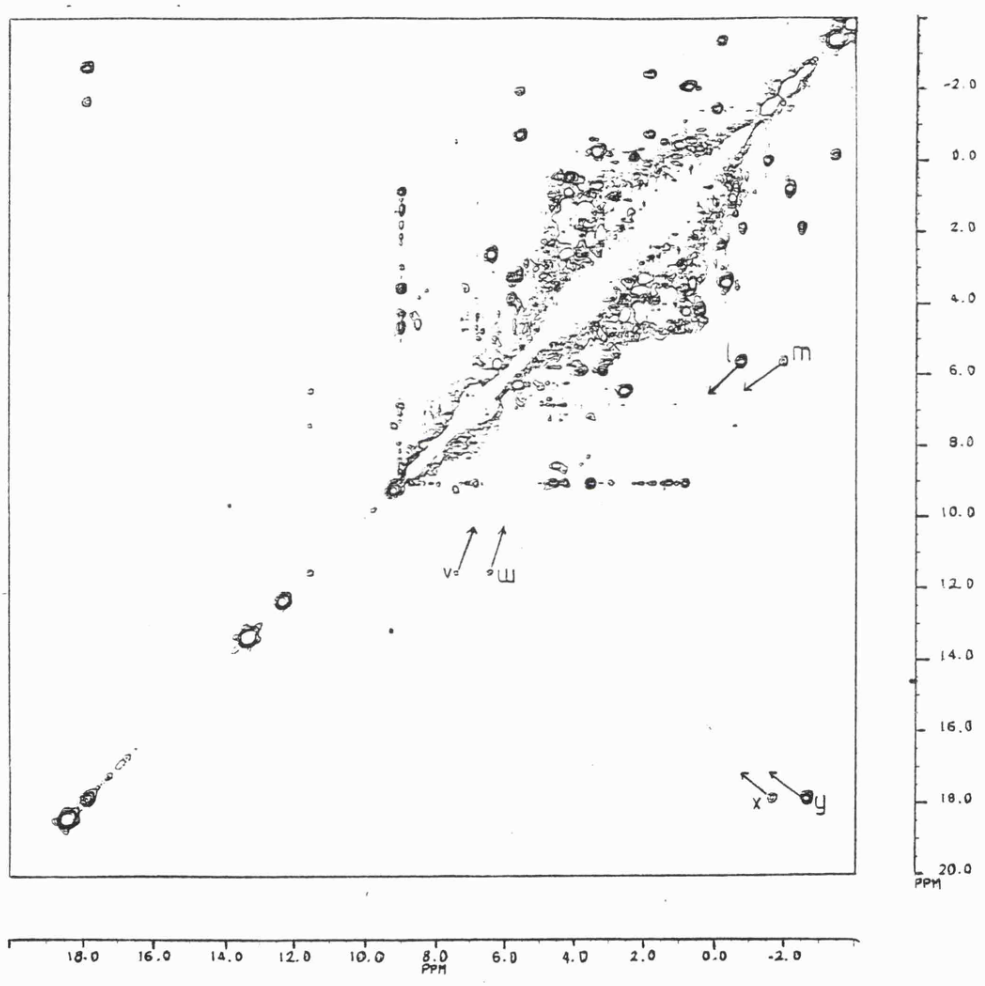


Figure 4.x.

$^{1}H^{1}H$ COSY spectrum of Equine Skeletal Cyanomet-myoglobin @27°C showing changes in chemical shift positions on going to 52°C

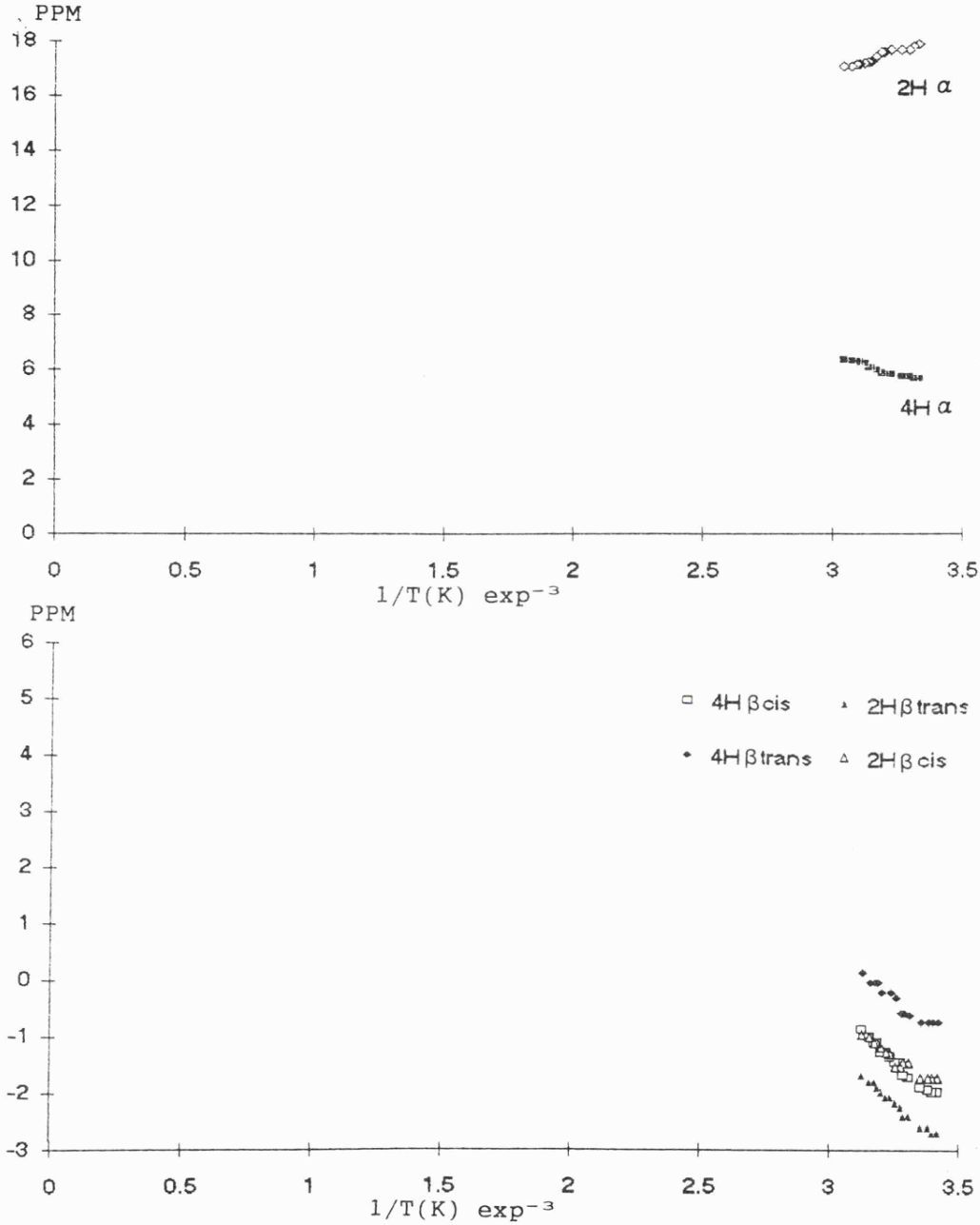


Fig. 4.xi

Proton Chemical Shifts of Haem 2-vinyl and proposed Haem 4-vinyl Resonances of Equine Cyanometmyoglobin Plotted Against Inverse Absolute Temperature

4.3 Assignments of Amino Acid Resonances.

4.3.1 Proximal Histidine Residue

The proximal histidine residue (His F8), which provides the axial ligand to the iron centre is common to all oxygen binding haemoproteins, and is one of only five amino acid residues conserved in all natural genetic variants of myoglobin.[150]

The interaction between the proximal histidine and the haem is pivotal in terms of the differential reactivity of haemoproteins towards oxygen and carbon monoxide[93], as the proximal histidine is found to move towards the haem plane on ligand binding.[150]

The resonances due to the proximal histidine residue should therefore be expected to experience considerable hyperfine shifts, due to their intimate approach to the haem.

The proton chemical shift of the $C_\beta H'$ resonance of the proximal histidine residue of cyanometmyoglobin has been assigned at ≈ 11.5 ppm[135], compared with a diamagnetic shift in carbon monoxymyoglobin of 1.55 ppm[149], and this resonance is identified as H29 in the one dimensional proton spectrum (Fig. 4.iv).

The proton proton COSY spectrum (Fig 4.ix) shows the presence of a pair of cross-peaks (identified as v and w) linking the resonance designated H29 with two further

resonances at ≈ 7.3 and ≈ 6.2 ppm respectively.

From the structure of the histidine residue (Fig. 4.xii) it can be seen that cross peaks would be expected to be manifested between the $C_\beta H'$ and both the $C_\beta H$ and $C_\alpha H$ resonances

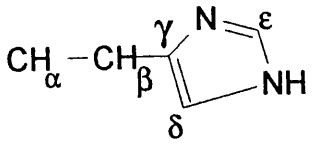


Fig. 4.xii

Structure of a Histidine Residue

thus leading to the assignment of the resonances at 6.2 ppm and 7.3 ppm as His F8 $C_\alpha H$ and His F8 $C_\beta H$, or vice versa.

Reference to Fig. 4.x and Table 4.iv shows that these resonances display strong temperature dependencies as predicted by their large hyperfine shifts, and Curie plots of chemical shift versus reciprocal absolute temperature should be expected to tend towards the corresponding diamagnetic values, which are given by Dalvit and Wright[149] as :-

$C_\alpha H \approx 2.90$ppm; $C_\beta H \approx 1.72$ppm; $C_\beta H' \approx 1.55$ppm

The Curie plots of these resonances are given in Figure 4.xiii, and by comparison of the known diamagnetic shifts with the observed intercepts I assign the resonances as follows :-

His F8 $C_\alpha H$ = 7.3ppm; His F8 $C_\beta H$ = 6.3ppm (@ 310K)[151]

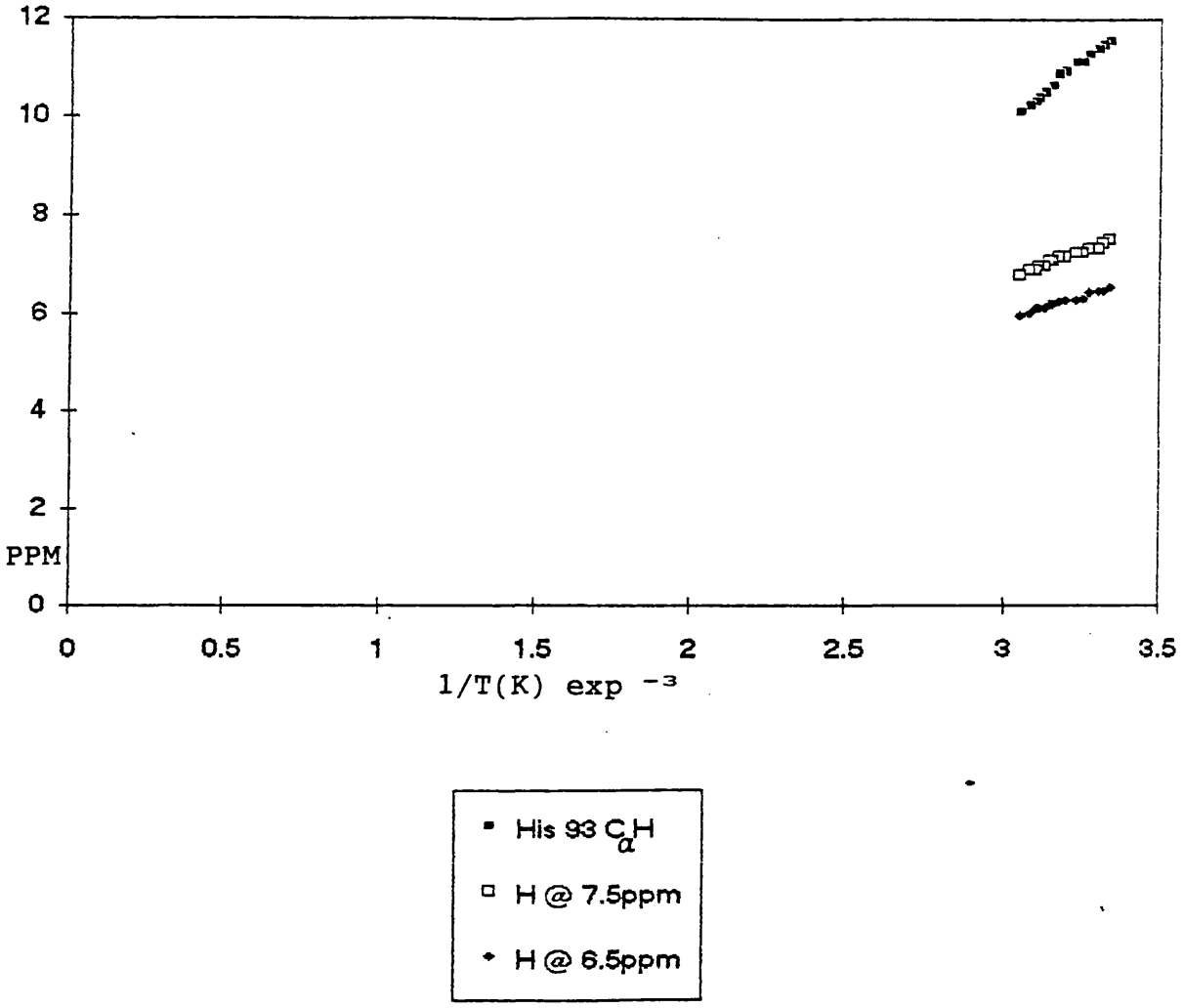


Fig. 4.xiii

Proton Chemical Shifts of the Proximal Histidine CβH'Resonance and Proposed CαH and CβH Resonances of Equine Cyanometmyoglobin, Plotted Against Inverse Absolute Experimental Temperature.

4.3.2 Isoleucine FG5

Isoleucine residue FG5 is located on the proximal side of the haem, over pyrrole ring II, with the majority of the residue being located within 7Å of the iron centre[85,150]. It is believed to play an important role in the control of ligand binding[82], and the equivalent residue, Valine FG5, in haemoglobin is implicated in inter-subunit interactions, with a variable steric interaction between this residue and the haem 4-vinyl group being proposed to be the mechanism, on a molecular level, of haemoglobin cooperativity. Consequently comprehensive assignments for this important residue would be a valuable tool in the future NMR studies of haemoproteins.

Proton assignments for this isoleucine residue, which gives rise to some of the farthest out-lying signals in the proton spectrum, have been published, and are given in Table 2.iii, and certain of these signals may be identified in the proton spectrum in Fig. 4.iv as :-

$$C_{\delta}H_3 = H2;\ C_{\gamma 2}H_3 = H3;\ C_{\gamma 1}H' = H6;\ C_{\gamma 1}H = H1$$

and these assignments are confirmed by the presence of the expected cross-peaks, designated :-

$$b = H2 \leftrightarrow H6 = \text{Ile FG5 } C_{\delta}H_3 \leftrightarrow \text{Ile FG5 } C_{\gamma 1}H'$$

$$c = H3 \leftrightarrow H\ @\ \sim 0.05\text{ppm} = \text{Ile FG5 } C_{\gamma 2}H_3 \leftrightarrow \text{Ile FG5 } C_{\beta}H$$

in the proton proton COSY spectrum (Fig. 4.ix). **NB** the H1 resonance in the one dimensional proton spectrum is off the scale of the proton proton COSY spectrum as presented.

The proton carbon-13 shift correlation spectrum (Fig. 4.vi) now clearly shows cross-peaks due to the two isoleucine FG5 methyl groups, correlating clearly to the carbon resonances designated C7 and C13 in the one dimensional carbon-13 spectrum (Fig. 4.v). The fact that these two resonances appear,in isolation, in such an unique spectral window should make them ideal NMR probes of the haem ligand binding site, and I assign them as follows :-

Ile FG5 $C_{\delta}H_3$ = C7 = 4.2ppm

Ile FG5 $C_{\gamma 1}H_3$ = C13 = 12.4ppm

(@ 310K)

whilst a further, more tentative assignment gives :-

Ile FG5 $C_{\gamma}H_2$ = C @ 16.4ppm.

4.3.3 Valine E11

Valine residue E11 is situated on the distal side of the haem, between pyrroles I and IV, with its point of closest approach to the iron centre being the γ_2 methyl group, situated some 4.14Å from the paramagnetic iron centre.[150] Valine E11 is believed to be implicated in the control of ligand binding.[82]

Surprisingly, however, considering its functional importance, and the high degree of paramagnetic shift to which it must be subject, only one assignment had been put forward for this residue, when this work was undertaken namely that of the α proton :-

Val E11 $C_\alpha H \approx -2.61$ppm[109,152]

This resonance is identifiable as H5 in the one dimensional proton spectrum (Fig. 4.iv) and reference to Table 4.ii demonstrates that this resonance displays a strongly temperature dependent chemical shift, as would be expected for a species in such close proximity to the haem.

The structure of a Valine residue is given in Figure 4.xiv :-

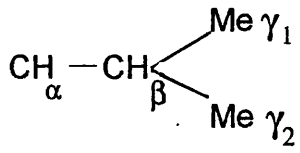


Fig. 4.xiv

The Structure of a Valine Residue

The cross-peak designated "f" in the proton proton COSY spectrum (Fig. 4.ix) shows the presence of a coupling between H5, the Val E11 $C_\alpha H$, and a resonance at ~ 1.9ppm, indicating

that this resonance is due to the Valine E11 β proton. A further cross-peak, designated "g" is manifested between the resonance at 1.9ppm and a resonance at -0.68ppm (which can be identified as H10 in the proton spectrum, Fig 4.iv), and reference to the structure of the valine residue (Fig 4.xiv) shows that this resonance could, therefore, be due to either to the γ_1 or to the γ_2 methyl group. The equivalent coupling to the second γ methyl group was not observed and was assumed to be obscured by the diagonal, even in those experiments in which the intensity of the diagonal was suppressed (COSY 45, COSY LR - see chapter 3). Since the Valine E11 γ_2 methyl group is situated closer to the paramagnetic iron centre than is the γ_1 methyl group[150], it might be expected to experience a greater paramagnetic contribution to its proton chemical shift, and consequently be resolved from the diagonal in the proton proton COSY spectrum, whilst the γ_1 methyl group is obscured. Consequently I make the following assignments :-

H @ 1.6ppm = Val E11 $C_\beta H$

H @ - 0.68ppm (H10) = Val E11 $C_{\gamma 2}H_3$ (or less likely $C_{\gamma 1}H_3$)

Furthermore this methyl group can be clearly identified in the proton carbon-13 shift correlation spectrum, and consequently the carbon-13 assignment may be made as follows :-

^{13}C @ 15.7ppm = Val E11 $C_{\gamma 2}H_3$ (@ 303K)

4.3.4 Other Isoleucine Residues

The amino acid sequence of equine skeletal myoglobin (Table 2.ii) reveals the presence of nine isoleucine residues, including that of isoleucine FG5 discussed in section 4.3.2. These nine isoleucines are as follows :-

B2,B11,E18,FG5,G2,G8,G12,G13,H18

The δ methyl carbons of isoleucine residues have been shown to appear in a characteristic region of the carbon-13 spectra of proteins, normally between approximately 11 and 12.5ppm[26], and a series of discrete resonances are clearly visible in this region in both the carbon-13 (Fig. 4.v) and the proton carbon-13 shift correlation (Fig 4.vi) spectra. The δ methyl carbons of eight of the nine sperm whale myoglobin isoleucine residues have been assigned[15], and these assignments are given below :-

Ile B9 @ 14.16ppm, Ile B11 @ 13.17ppm, Ile G13 @ 12.85ppm, Ile G8 @ 12.75ppm, Ile G12 @ 12.51ppm, Ile H18 @ 11.40ppm, Ile G2 @ 9.80ppm, Ile E18 @ 9.20ppm.

Since isoleucine B9 is replaced by a valine residue on going from sperm whale to equine myoglobin[121], the remaining sharp carbon resonance in this region, C10 @ 11.22ppm (300K) must be due, instead, to the δ methyl carbon of the only isoleucine residue found in equine myoglobin, which is absent in the sperm whale variant, namely isoleucine B2. I therefore assign the carbon resonance designated C10 as :-

C10 (@11.22ppm, 300K) = Ile B9 $C_{\delta}H_3$

Proton chemical shifts may be correlated to these carbon shifts from the $^{1}H^{13}C$ shift correlation spectrum (Fig. 4.v) as follows :-

C8 = H @ 0.45ppm, C9 = H @ 1.00ppm, C10 = H @ 1.45ppm,
C11 = H @ 1.60ppm, C12 = H @ 0.68ppm, C13 = H @ -3.05ppm,
C14 = H @ 0.70ppm, C15 = H @ 1.10ppm

Now C13 can be clearly seen to be equivalent to H3 @ -3.05ppm (310K), and this spectroscopic unit {C13-H3} has already been unambiguously assigned to the γ methyl group of the isoleucine FG5 residue (see section 4.3.2 above) and therefore it is clear that the published assignment of the isoleucine G8 δ methyl carbon[115] is in error.

Examination of the crystal structure data reveals that isoleucine G8 is in sufficiently close approach to the haem pocket to experience some degree of paramagnetic shift, and it is therefore not surprising that it does not occur within the same range as the other isoleucine δ methyls. I therefore tentatively assign the Ile G8 δ Me as the resonance designated C16, being the next most upfield shifted carbon resonance, with chemical shifts

^{13}C = 15.1ppm, ^{1}H = 0.8ppm (310K)

summarising this data gives the following assignments :-

Ile E18 δ Me = H @ 0.45ppm (310K)		
Ile G2 δ Me = H @ 1.00ppm	"	
Ile B2 δ Me = H @ 1.45ppm	"	C @ 11.22ppm (300K)
Ile H18 δ Me = H @ 1.60ppm	"	
Ile G12 δ Me = H @ 0.68ppm	"	
Ile G13 δ Me = H @ 0.70ppm	"	
Ile B11 δ Me = H @ 1.10ppm	"	
Ile G8 δ Me = H @ 0.80ppm	"	C @ 15.10ppm (300K)

4.4 Multiplicity Determination of Signals in the Heteronuclear Shift Correlation Experiment.

The aliphatic region of the $^{1}H^{13}C$ shift correlation spectrum is extremely complex, indeed reference to the amino acid sequence of equine skeletal myoglobin reveals the presence of ninety amino acid side chain methyl groups alone. A methodology was therefore sought, whereby this crowded spectral region might be simplified and resolved.

The efficacy of combining the data obtained from carbon-13 spectra, edited according to the number of attached protons[144], with the large temperature dependence of carbon-13 chemical shifts, in order to resolve and identify many of the methyl groups within a protein spectrum has already been demonstrated[61], using horse cytochrome c, and this technique

has here been extended to the larger myoglobin molecule. (See Fig. 4.vii) Additionally, the pulse sequence employed in the heteronuclear shift correlation experiment is related to the "INEPT" (insensitive nuclei enhanced by polarisation transfer) method of Morris and Freeman[146], and consequently shares the advantage with this method, that subspectra may be generated, by means of a variable length delay within the pulse sequence, such that spectral units of different multiplicities experience differential enhancements. (See Chapter 3)

Specifically, by selection of appropriate delays, proton carbon-13 shift correlation spectra may be recorded in which methine signals are optimised at the expense of both methyl and methylene signals, or in which methyl and methine groups are enhanced and positive, whilst methylene groups are enhanced and negative. An example of this latter type of optimised spectrum is shown in Figure 4.xv.

Combining these techniques the following conclusions may be drawn, and these are summarised in Figure 4.xvi.

All signals in the $^{1}H^{13}C$ shift correlation spectrum (Figs. 4.vi, 4.xvi) appearing below 20ppm in the carbon-13 dimension, and/or below -1ppm in the proton dimension are due to methyl groups, whilst in the region above 40ppm in the carbon-13 dimension the signals are predominantly due to methine groups, together with a few methylene groups. In the most crowded region of the spectrum, between 20 and 40ppm in the carbon-13 dimension a complex combination of around sixty methyl, methylene, and methine groups are found, however the

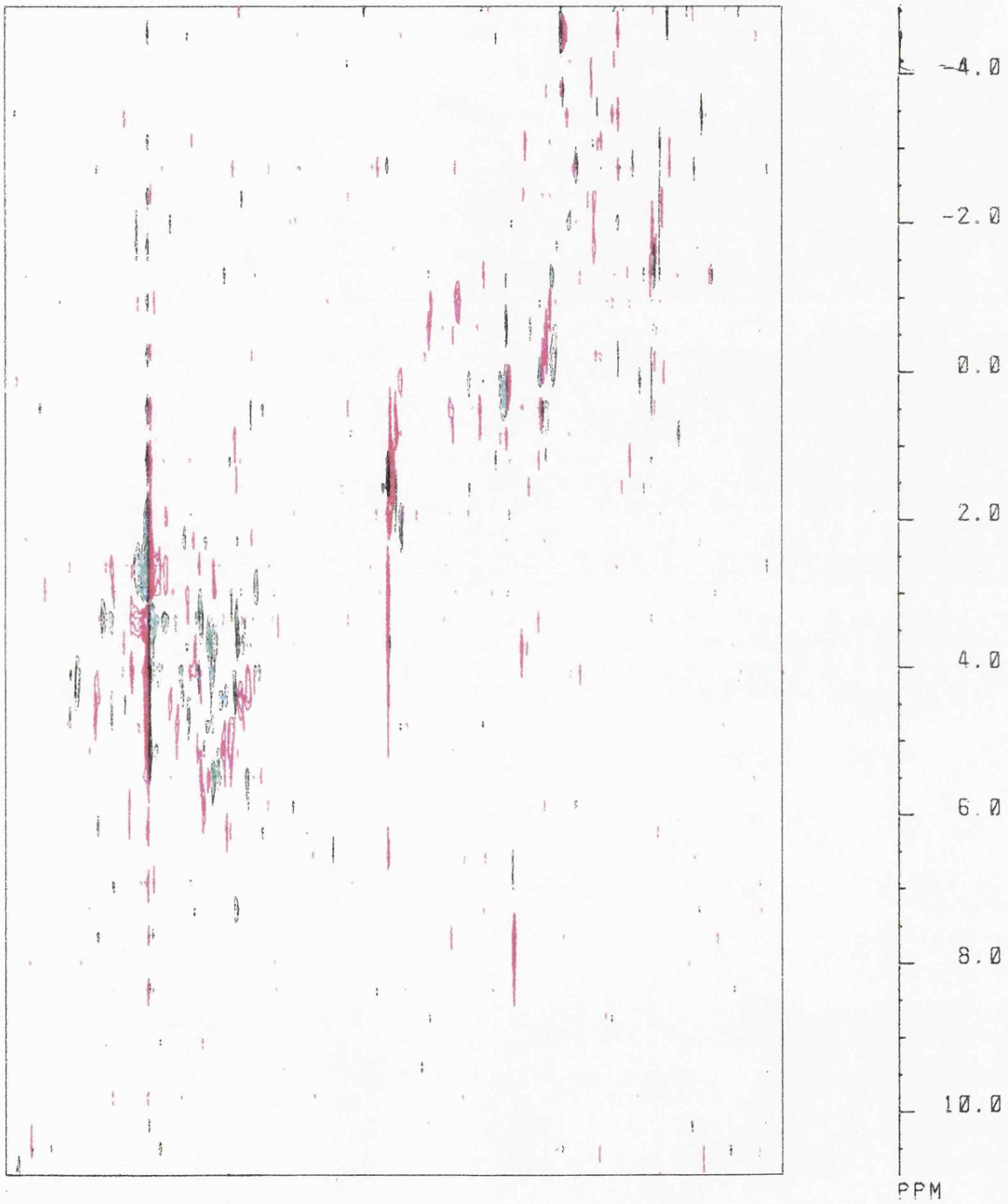


Fig 4.xv

$^{1}H^{13}C$ Shift Correlation Spectrum Optimised to Give CH_3 and CH Signals Positive (Red), and CH_2 Signals Negative (Green)

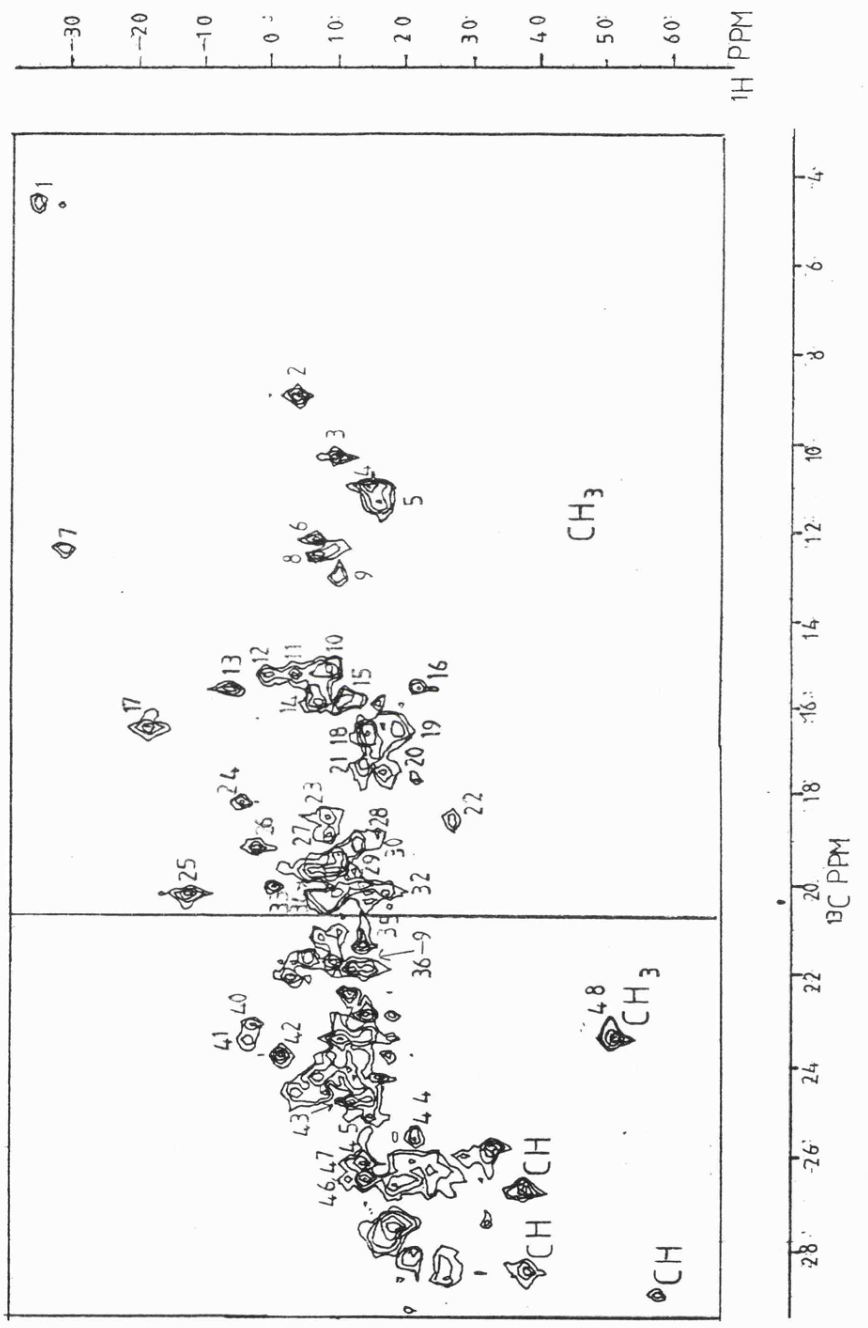


Fig 4.xvi

Section of the $^{1}H^{13}C$ Shift Correlation Spectrum of Equine Cyanometmyoglobin, Annotated to Show the Multiplicities of Certain Signals.

use of edited spectra suggests that between 25 and 30ppm in the carbon-13 dimension only five or six methyl groups are found, and no methyls are located above 30ppm.

The pair of strong signals centred on 40ppm (^{13}C) as well as the nearby small peak arise from methylene groups, whilst in the region delineated by 20-30ppm (^{13}C) and 3.5-6ppm (^{1}H) four signals are located, of which the signal with a carbon-13 chemical shift of 23.2ppm, labelled 48 in Fig 5.xvi, is due to a methyl group, the three remaining signals in this region arising from methine groups.

Altogether forty eight out of the possible ninety methyl groups have been unambiguously identified using this method, and these are identified in Fig 4.xvi, and their proton and carbon-13 chemical shifts (measured at 303K) are presented in Table 4.v.

4.5 Summary of Results

Many significant NMR assignments for myoglobin are presented in this work, and these are summarised in Table 4.vi. Notably, the full Proton and carbon-13 assignments for all four Haem methyl groups are presented, resolving the previous ambiguity surrounding the assignment of the haem 3-methyl group.

Peak	^{1}H shift	^{13}C shift	Peak	^{1}H shift	^{13}C shift
1	-3.5	4.5	24	-0.5	17.9
2	0.45	8.8	25	-1.3	20.0
3	1.0	10.2	26	-0.2	18.9
4	1.4	10.7	27	0.8	18.6
5	1.6	11.2	28	1.2	18.8
6	0.65	12.0	29	0.9	19.0
7	-3.1	12.2	30	0.8	19.5
8	0.7	12.4	31	0.0	19.8
9	1.0	12.8	32	1.7	20.0
10	0.9	14.9	33	1.2	19.9
11	0.3	15.1	34	0.9	19.9
12	-0.1	15.1	35	1.3	21.0
13	-0.6	15.4	36	0.7	20.0
14	0.7	15.6	37	0.3	21.3
15	0.95	15.9	38	0.9	21.7
16	2.2	15.3	39	1.2	21.8
17	-1.9	16.2	40	-0.3	22.8
18	1.4	16.3	41	-0.35	23.2
19	1.9	16.1	42	0.15	23.5
20	2.1	17.1	43	1.2	24.5
21	1.3	16.9	44	1.5	24.9
22	2.6	18.3	45	2.15	25.4
23	0.8	18.2	46	1.4	25.9
24	-0.5	17.9	47	1.4	26.3
			48	5.2	23.2

Table 4.v

Chemical Shifts of Peaks Identified as Methyls in the $^{1}H^{13}C$ Shift Correlation Spectrum of Equine Cyanometmyoglobin at 303K

Assignment	^{1}H shift (ppm)	^{13}C shift (ppm)
Haem 3-methyl	5.09[a]	-13.64[a]
Haem 7-CH α'	-1.7[a]	-28.0[a]
Haem 4-CαH	5.8[b]	
Haem 4-CβH_{trans}	-0.6[b]	
Haem 4-CβH_{cis}	-1.8[b]	
His F8 CαH	7.3[a]	
His F8 CβH	6.3[a]	
Ile FG5 CδH_3		4.2[a]
Ile FG5 C$\gamma_1 H_3$		12.4[a]
Ile FG5 CγH_2		16.4[a]
Val E11 CβH	1.6[a]	
Val E11 C$\gamma_2 H_3$	-0.68[a]	15.7[b]
Ile E18 CδH_3	0.45[a]	
Ile G2 CδH_3	1.00[a]	-
Ile B2 CδH_3	1.45[a]	11.22[c]
Ile H18 CδH_3	1.60[a]	
Ile G12 CδH_3	0.68[a]	
Ile G13 CδH_3	0.70[a]	
Ile B11 CδH_3	0.80[a]	15.10[c]

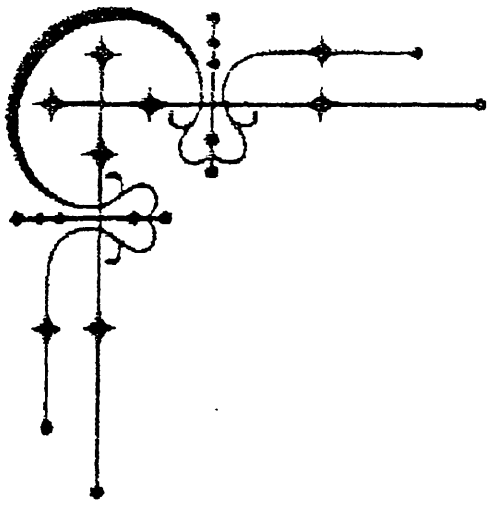
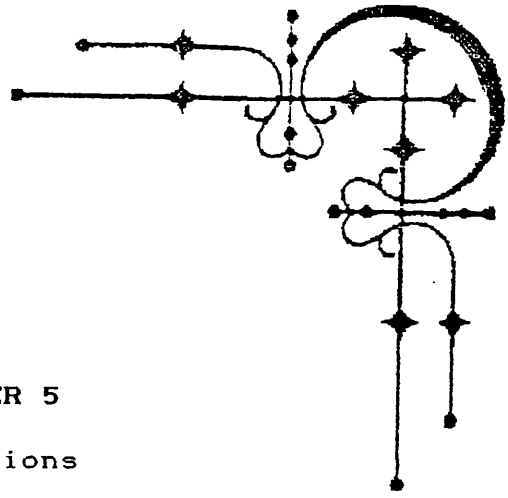
a - shifts measured @ 310K
b - shifts measured @ 303K
c - shifts measured @ 300K

Table 4.vi

Summary of assignments referred to in Text

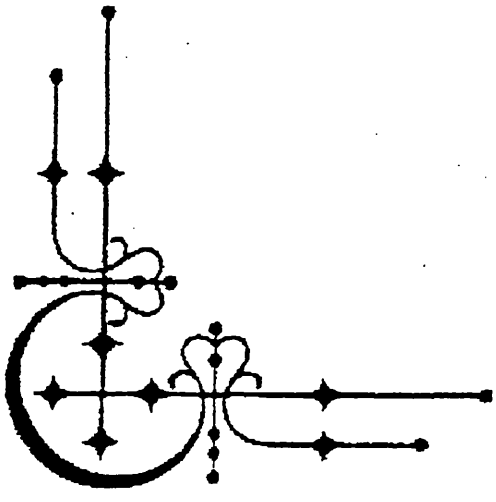
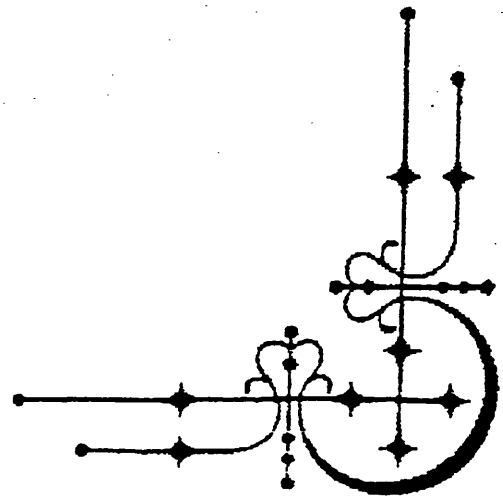
Additionally full proton assignments are presented for the important haem 4-vinyl group, which experiences a smaller degree of paramagnetic shift than does the 2-vinyl group, and which, as a consequence had been previously unassigned in the proton spectrum, and an assignment to the haem 7-propionate group is proposed.

Assignments in both the proton and carbon-13 spectra are also presented for several of the most functionally relevant amino acid residues of the molecule, including the ubiquitous and physiologically vital proximal histidine residue, which plays an important role in modulating the activity of the haem iron centre, and the isoleucine FG5 , and valine E11 residues, both of which are located within the haem pocket of the molecule, and are believed to be intimately involved in protein function.[82]

CHAPTER 5

Conclusions

The assignments presented in this research, should prove a valuable basis for future studies on the structure/function relationships not only of equine myoglobin, but also of the many other closely related haemoproteins for which it can be used as a model.

The research has been of value, however, not merely in terms of the information gleaned on myoglobin itself, but equally as a prototype study for the application of certain of the NMR techniques employed herein; which had been previously believed to be inapplicable to the study of paramagnetic species, or which had previously been employed solely in the study of much smaller molecules; to other less well characterised proteins.

Two-dimensional shift correlated spectroscopy (COSY), for example, was originally believed to be inapplicable to paramagnetic species since the line broadening and rapid loss of coherence caused by paramagnetic-induced relaxation was thought to interfere with COSY peak detection.

By using proton-proton COSY spectra to make certain of the assignments described above, as well as by utilising the manifestation of COSY cross-peaks in spectral windows unique to the paramagnetic form of myoglobin, in order to characterise these resonances, this research has demonstrated that the COSY technique may indeed be profitably applied to the hyperfine shifted resonances of paramagnetic haemoproteins. (See also Appendix A which details evidence of similar conclusions published by other workers in 1990/1991)

The usefulness of spectroscopic techniques rarely used in the study of any biological macromolecule; such as two-dimensional heteronuclear shift correlation spectroscopy, and of both one and two-dimensional techniques involving the editing of the natural abundance carbon-13 spectrum according to the number of attached protons; to the study of paramagnetic haemoproteins has also been demonstrated.

APPENDIX A

More Recent Developments in the NMR of Proteins

"The Road goes ever on and on,
Down from the door where it began,
Now far ahead the road has gone,
And I must follow if I can,
Pursuing it with eager feet,
Until it meets some larger way,
Where many paths and errands meet,
And wither then I cannot say."

J.R.R. Tolkein
The Lord of the Rings

The study of proteins and other biological macromolecules by NMR spectroscopy continues to be a field of interest. Recently the solution structures of many proteins have been assigned and solved using the Wüthrich method of sequential assignment, as described in Chapter 1. For example The three dimensional structure of the small protein Complement Factor 5a, a pharmacologically interesting molecule involved in inflammation, for which no crystal data is available, has been solved by NMR[158], and the proton spectrum of human insulin has been assigned[159].

The predicted growth of interest in three and four dimensional techniques has led to a wealth of experimental strategies. Recent three dimensional methodologies for obtaining amino acid side chain assignments in biological macromolecules include those that begin, not by establishing ^{1}H-^{1}H J couplings as did earlier methods[160,161], with the resulting disadvantage of poor resolution, but by the transfer of magnetisation from ^{1}H to ^{13}C which results in an increase in sensitivity.[162,163]

The HCCH approach[162] for example follows this initial transfer of magnetisation with a COSY type mixing pulse during which magnetisation is transferred between ^{13}C coupling partners, before transferring back to attached protons via a reversed INEPT sequence. Three dimensional techniques of this type have been used in the solution of the structure of the protein interleukin-1β.[162,164]

Continued research into computational methods of converting NOE-derived distance constraints into structures has led to the re-examination and improvement of the main chain directed algorithm described in Chapter 1[165,166], and the use of the updated version has been demonstrated with the protein ubiquitin[166].

Work has been carried out which analyses proton chemical shift data in proteins of known crystallographic structure. "Structural" chemical shifts which depend only on local conformation have been derived for several proteins, by means of subtracting from their observed shifts the shifts of equivalent protons in short random coil polypeptides, ring current shifts[167,168], and contributions from peptide groups[169], and useful structural correlations have been found. For example it has been shown that the Cα protons of amino acid residues located within α-helices are predominantly shifted upfield, whilst those located in β-pleated sheets experience largely downfield shifts.[168,170]

Isotopic labelling strategies permit NMR assignments to be made in increasingly large proteins, for example by employing ^{2}H labelling, about 50% of the assignments for the 25kD TRP repressor protein have been obtained.[171] and Le Master has recently reviewed the use of deuterium enrichment in the NMR of biological macromolecules[172].

Both one and two dimensional ^{13}C-^{1}H heteronuclear multiple quantum coherence experiments have been carried out on the enzyme chloramphenicol acetyltransferase, into which

histidine residues, specifically labelled with ^{13}C at the ring 2 carbon had been incorporated, and six of the C_2H resonances have been clearly identified, demonstrating that selective isotopic enrichment and multiple quantum coherence techniques can be used to identify individual proton NMR signals in a protein with molecular mass of 75000.[173]

Both selective and uniform labelling of proteins with nitrogen-15, and ^{15}N-^{13}C double enrichment strategies have been employed. The DEPT sequence[144] has been employed to enhance the sensitivity of ^{15}N NMR spectra of uniformly ^{15}N labelled proteins[xvi], and by spectral editing the ^{15}N spectra of these proteins have been decomposed into subspectra, according to the number of attached protons.[174,175] Additionally the assignment of ^{15}N resonances followed by two dimensional heteronuclear shift correlation experiments has allowed the amide proton resonance assignments of proteins for which the conventional sequential assignment sequence had proved ineffective.[174]

The binding sites of large proteins have also been studied, without the need to obtain assignments for the entire protein by "editing" the NMR spectrum using ^{13}C or ^{15}N labelled ligands or inhibitors.[158,176]

The past two years have also seen several developments specifically related to myoglobin.

Oxygen-17 and carbon-13 NMR studies have been carried out on several $C^{17}O$ and ^{13}CO ligated haemoproteins including human and rabbit haemoglobins, horseradish peroxidase and wild type sperm whale myoglobin and distal point mutants thereof,[177] and

the ^{17}O and ^{13}C chemical shifts of the bound carbon monoxide have been determined, showing modulation of the nature of the Fe-C bond by different distal residues.[177] Solid-state oxygen-17 NMR studies have allowed the ^{17}O NMR spectra of oxyhaemoglobin and oxymyoglobin to be recorded for the first time [178] providing a direct probe of metal-oxygen interactions.

Recently the carbon-13 NMR resonances of the proximal histidyl imidazole ring carbons of horse deoxymyoglobin and deoxyhaemoglobin have been reported,[179] and the sensitivity of the imidazole ring ^{13}C resonances to the iron-histidine bond demonstrated.

NMR studies into the intercalation of the haem group into its pocket have continued. Three types of interaction hold the haem in place within the protein, namely the iron-histidine bond, interactions between the haem π system and hydrophobic groups in the haem pocket, and electrostatic interactions between the haem propionate side chains and residues bordering the haem cavity, as described in Chapter 2. Lecomte and Cocco have carried out NMR investigations into des-iron myoglobin,[180] and NOE data suggest that the structure of this complex is very similar to that of the native holoprotein, with the protoporphyrin bound at the same height as the haem, and with the same contacts with residues lining the cavity, implying that the positioning of the porphyrin is governed solely by hydrophobic and electrostatic interactions. In addition the distal side of the haem pocket appears largely structurally unperturbed whilst the His-F8 environment remains largely

intact. Apomyoglobin is also under investigation to see how this will compare structurally, and preliminary studies have been published.[181] The alternative 180° insertion isomers, however, are not observed in the des-iron complex, even in freshly prepared solutions, and it is believed that the complex equilibrates too rapidly, in the absence of the iron-His F8 bond, for this to be observed.[180]

Similarly NMR studies by Deeb and Peyton on the diamagnetic tin-porphyrin myoglobin have shown almost identical NOE patterns for both tin-myoglobin and carbonmonoxymyoglobin, whilst ^{119}Sn NMR evidences the presence of His F8-metal coordination, suggesting that the porphyrin maintains the same orientation in both complexes, with very little apparent deformation of the binding site.[182] Furthermore proton and tin-119 NMR studies of the interaction of the tin-protoporphyrin with apomyoglobin have demonstrated the existence of a long lived intermediate, possibly the 180° insertion isomer, which would support the idea of rapidly exchanging porphyrin insertion isomers in the absence of a His F8-metal bond.[183]

Myoglobins reconstituted with various modified hemins, bearing 3,[184] or 4[185] propionate groups (with the remainder of the haem sites being occupied by methyl groups) have also been studied. The orientations of these modified hemins within the protein matrix were monitored by relating the positions of the haem methyls to various characteristically hyperfine shifted haem pocket amino acid residues, by NOE methods. It has been shown that the haem pocket can accommodate a propionate at

virtually any non-native site in the protein matrix, and characteristic hyperfine shifts suggest little perturbation of the local pocket structure, although in addition to the native propionate sites a particular affinity is shown for the native two-vinyl position, where the propionate is thought to occupy the "xenon hole". Additionally the rotational "hopping" of the haem about the iron-His F8 bond, previously described for des-propionate myoglobins[111,112], has been shown to occur in these complexes,[h] and is now believed to be a general property of the haem pocket.

The trimethyl phosphine derivative of sperm whale myoglobin has been studied, and its spectrum compared with that of carbonmonoxymyoglobin[186] and the greater steric hindrance in the trimethyl phosphine derivative has been shown to give rise to differences in conformation in the haem pocket, especially affecting those distal residues closest to the haem plane.

Proton NMR has been employed to demonstrate changes in the characteristic haem methyl chemical shifts, on going from wild type sperm whale myoglobin to point mutants with the His-F8 residue replaced by Cys or Tyr.[187] These haem methyl shift patterns were compared with those of haem peroxidase enzymes, showing that the electronic structure of the mutants resembles that of the corresponding haem enzymes rather than that of the parent myoglobin. This supports the theory that ligands such as phenolate and thiolate are functionally relevant to catalytic activity, electron donation to the iron facilitating heterolytic O-O bond cleavage.[187]

The chemical shifts of the haem methyl groups of sperm whale cyanometmyoglobin have been used to map out the paramagnetic metal centred dipolar field, using a computer algorithm based on the mathematical relationship between the metal centred pseudocontact shift and the spatial coordinates of the nucleus with respect to the haem.[188]

Both the proton and the carbon-13 assignments of the haem methyl groups,[189] and haem vinyl groups,[190] of shark cyanometmyoglobin have been published, and the previously published ^{1}H and ^{13}C assignments for the haem methyl groups of sperm whale cyanometmyoglobin[134] have been revised,[189] the newly published assignments corresponding with those presented in this work.

The hyperfine shift pattern of shark cyanometmyoglobin has been shown to closely resemble that of sperm whale cyanometmyoglobin, with the haem methyls displaying hyperfine shifts intermediate to those of sperm whale and Aplysia limacina cyanometmyoglobins.[189]

The in plane asymmetry of the electronic structure of the porphyrin ring has been shown to be dependent upon the angle between the projection of the proximal histidyl imidazole plane onto the haem plane, and the N_{II}-Fe-N_{IV} vector $(\phi)o$, and consequently this intermediate spread of haem methyl hyperfine shifts has been interpreted as projecting a value of ϕ for shark myoglobin as being between 19° and 29°, being the values reported for sperm whale and Aplysia limacina myoglobins respectively.[189]

The orientations of the haem vinyl groups of sperm whale and shark cyanometmyoglobins with respect to the π system of the porphyrin have been inferred from their hyperfine shifts,[189] using the computational method discussed above.[191] These orientations are of interest since they directly modulate the electron withdrawing ability of the vinyl groups, and therefore the electronic structure of the haem, and thus its reactivity are consequent upon them.

It has been demonstrated that two-dimensional nuclear Overhauser effect spectroscopy can be applied to paramagnetic haemoproteins by using the X-ray crystal coordinates to interpret cross-peaks,[192,193] and subsequently the two-dimensional bond correlation methods of COSY, double quantum filtered COSY, and two-dimensional homonuclear Hartmann-Hahn spectroscopy, have been successfully applied to sperm whale cyanometmyoglobin,[194] confirming the conclusions, presented in this report, that these methods are indeed applicable to paramagnetic metalloproteins.

As a consequence of these new two-dimensional NMR studies several assignments for sperm whale cyanometmyoglobin have been made,[192,193,194] certain of which independently confirm assignments made in this work, and these are presented in Table A.i.

The applicability of the techniques of $^1H^1H$ COSY and NOESY to paramagnetic proteins has further been demonstrated by the publication of assignments derived solely from the tandem use of these two techniques, for the downfield resolved protons

Residue	Spin	Chem. Shift‡
Leu 29	$C_{\gamma}H$	3.96
	$C_{\delta 1}H_3$	3.81
	$C_{\delta 2}H_3$	5.53
Phe 33	$C_{\delta}H_3$	8.04
	$C_{e}Hs$	8.37
	$C_{\zeta}H$	8.40
Phe 43	$C_{\zeta}H$	17.27
Phe 45	$C_{\delta}Hs$	7.69
	$C_{e}Hs$	8.05
His 64	$C_{\beta}H$	3.81
	$C_{\beta}H'$	4.41
Thr 67	$C_{\alpha}H$	2.5
	$C_{\beta}H$	2.69
	$C_{\gamma}H_3$	-1.56
Ala 71	$C_{\alpha}H$	3.46
	$C_{\beta}H_3$	-0.12
Leu 89	$C_{\alpha}H$	8.3
	$C_{\beta 1}H$	4.4
	$C_{\beta 2}H$	3.7
	$C_{\gamma}H$	5.7
	$C_{\delta 1}H_3$	3.8
	$C_{\delta 2}H_3$	3.1
Ser 92	$O_{\gamma}H$	9.71
His 93	$C_{\alpha}H$	7.3
Ala 94	$C_{\alpha}H$	6.47
	$C_{\beta}H_3$	2.66
His 93	$C_{\beta}H$	1.32
	$C_{\beta}H'$	-0.48
	$C_{\delta}H$	11.07
Leu 104	$C_{\gamma}H$	0.07
	$C_{\delta 2}H_3$	-1.49
Ile 107	$C_{\gamma}H_3$	0.37

Table A.i

Some recently published assignments for myoglobin[192,193,194]

Haem substituent	Spin	Chem. shift‡
Haem meso Hs	α	4.4
	β	2.09
	γ	5.98
	δ	4.09
Haem 4-vinyl	$C_{\alpha}H$	5.7
	$C_{\beta}H_{cis}$	-1.5
	$C_{\beta}H_{trans}$	-0.4
Haem 6-propionate	$C_{\beta 1}H$	1.7
	$C_{\beta 2}H$	-0.4

Table A.i cont.

Some recently published assignments for myoglobin[192,193,194]

‡ Taken from Refs. 192,193 and 194 Measured at 35°C and pH 8.6

of equine met-azido myoglobin, a mixed spin state molecule, which displays significantly better spectral resolution than does cyanometmyoglobin, and which provides a good model system for certain physiological forms of haemoglobin.[195]

REFERENCES

1. Pollack A. New York Times (1988), 15/3/88, C1,C11 (quoted in 20).

2. Wüthrich K. Nato ASI Ser A (1989), 183, 69-78.

3. Wagner G. *et al.* J.Mol.Biol. (1987), 196, 611-39.

4. Shultze P. *et al.* J.Mol.Biol. (1988), 203, 251-68

5. Furey W.F. *et al.* Science (1986), 231, 704-10

6. La Mar G.N. *et al.* J.Mol.Biol. (1983), 168, 887-96

7. Cozzone P. *et al.* Proc. Nat. Acad. Sci. USA. (1975), 72 2095-8.

8. Bemski G. *et al.* Acta Cient. Venezolana (1980), 31, 125-7

9. Howarth O.W. and Lian L.Y. J.C.S.Chem.Comm. 1983, 434

10. Kopple K.D. *et al.* J.Am.Chem.Soc. (1969), 91, 4264

11. Oldfield E. *et al.* J.Biol.Chem. (1975), 250, 6381-402

12. Wilbur D.J. and Allerhand A. J.Mol.Biol. (1976), 251, 5187-94

13. Carver J.A. and Bradbury J.H. Biochem. (1984), 23, 4890-905

14. Williams R.J.P. Eur.J.Biochem. (1984), 183, 479-97

15. Roberts G.C.K. Nato ASI Ser A (1989) 183, 69-78

16. Dobson C.M. Nato ASI Ser A (1989) 183, 193-208

17. Bloch F. *et al.* Phys. Rev. (1946), 69, 127

18. Purcell E.M. *et al.* Phys. Rev. (1946), 70, 986-7

19. Saunders M. *et al.* J.Am.Chem.Soc. (1957), 79, 3289-90

20. Wüthrich K. Acc.Chem.Res. (1989), 22, 36-44

21. Macdonald C.C. and Philips W.D. J.Am.Chem.Soc. (1967), 89,

6332-41

22. Macdonald C.C. and Philips W.D. J.Am.Chem.Soc. (1969), 91, 1513

23. Bak B. *et al.* J.Mol.Spect. (1968), 26, 78

24. Bradbury J.H. and Scheraga H.A. J.Am.Chem.Soc. (1966), 88, 4240-6

25. Gibbons W.A. *et al.* Nature (1970), 227, 840-2

26. Howarth O.W. and Lilley D.M.J. Prog.Nucl.Mag.Res.Spect. (1973), 12, 1

27. Conti F. and Paci M. FEBS. Lett. (1971), 17, 149-52

28. Oldfield E. and Allerhand A. Proc.Natl.Acad.Sci.USA, (1973), 70, 3531-5

29. Jones W.C. *et al.* J.Biol.Chem. (1977), 251, 7452-60

30. Schejter *et al.* J.Biol.Chem. (1978) 253, 3768-70

31. Redfield A.G. and Gupta R.K. Cold Spring Harbor Symp. Quant. Biol. (1971), 36, 405-11

32. Balaram P. *et al.* J.Am.Chem.Soc. (1972) 94 4015-7

33. Jeener J. Ampere International Summer School, Basko Polje, Yugoslavia (1971)

34. Gordon S.L. and Wüthrich K. J.Am.Chem.Soc. (1978), 100, 7094-6

35. Wagner G and Wüthrich K. J.Magn.Reson. (1979), 33, 675-80

36. Aue W.P. *et al.* J.Chem.Phys. (1976), 64, 2229

37. Ernst R.R. *et al.* "Principles of Nuclear Magnetic Resonances in One and Two Dimensions" (1987) Clarendon Press, Oxford.

38. Nagayama K. *et al.* Biochim.Biophys.Res.Commun. (1977), 78,

99-105

39. Williamson M.P. *et al.* J.Mol.Biol. (1985), 182, 295-315

40. Wüthrich K. "NMR of Proteins and Nucleic Acids" (1986) Wiley, New York.

41. Wagner G. and Wüthrich K. J.Mol.Biol. (1982), 155, 347-66.

42. Braun W. *et al.* Biochim.Biophys.Acta. (1981), 667, 377.

43. Havel T. and Wüthrich K. J.Mol.Biol. (1985), 182, 281

44. Beusen D.D. and Marshall G.R. Nato ASI Ser A (1989) 183, 97-109.

44a. Wüthrich K. Science (1989), 243, 45-50.

45. Englander S.W. and Wand A.J. Biochem. (1987), 26, 5953.

46. Englander S.W. *et al.* Tech. Protein Chem. (1989), 207-22.

47. Oh B.H. *et al.* J.Am.Chem.Soc. (1989), 111, 3083-5.

48. Bax A. *et al.* J.Am.Chem.Soc. (1989), 111, 408-0.

49. Stockman B.J. and Markley J.L. Nato ASI Ser A (1989) 183, 155-92.

50. Le-Master D.M. and Richards F.M. Biochem. (1988), 27, 142.

51. Crepsi H.L. *et al.* Science (1968), 161, 795.

52. Greisinger C. *et al. J.Magn.Reson. (1987), 73, 574-9.*

53. Greisinger C. et al. J.Magn.Reson. (1989), 84, 14-63.

54. Vuister G.W. and Boelens R. J.Magn.Reson. (1987), 74, 328.

55. Vuister G.W. et al. J.Magn.Reson. (1988), 80, 176-85.

56. Marion D *et al.* J.Am.Chem.Soc. (1989), 111, 1515-7.

57. Wider G. *et al.* J.Magn.Reson. (1989), 85, 426-31.

58. Chan T.M. and Markley J.L. J.Am.Chem.Soc. (1982), 104, 4010-11.

59. Chan T.M. and Markley J.L. Biochem. (1983), 22, 5996-6002.

60. Santos H. and Turner D.L. FEBS. Lett. (1987), 226, 179-185.

61. Santos H. and Turner D.L. FEBS. Lett. (1986), 194, 73-77.

62. Waterhous D.V. *et al*. J.Lipid Res. (1985), 26, 4667-9.

63. Bhacca N.S. *et al*. J.Am.Chem.Soc. (1983), 105, 2538-44.

64. Waterhouse A.L. *et al*. J.C.S. Perkin Trans. (1985), 1011.

65. Derome A.E. Nat. Prod. Rep. (1989), 6, 111-141.

66. Rinaldi P.L. and Sweicinski F.J. J.Magn.Reson. (1989), 82, 369-73.

67. Raftrey M.A. *et al*. Cold Spring Harbour Symp. Quant. Biol. (1971), 36, 541.

68. Müller N. and Mead R.J. Biochem. (1973), 12, 3831.

69. Gerig J.T. and Roe D.C. J.Am.Chem.Soc. (1974), 96, 233.

70. Gerothanassis I.P. *et al* Magn.Reson. Chem. (1985), 23, 659.

71. Baltzer L. J.Am.Chem.Soc. (1974), 109, 3479.

72. South T.L. *et al*. J.Am.Chem.Soc. (1989), 111, 395-6.

73. Dwek R.A. "Nuclear Magnetic Resonance in Biochemistry (Applications to Enzyme Systems) (1973), Clarendon Press, Oxford.

74. Mayer A. *et al*. J.Mol.Biol. (1974), 86, 749-56.

75. Wüthrich K. Struct. Bonding (1970), 8, 53-121.

76. Shulman R.G. *et al*. Probes Struct. Funct. Macromol. Membranes Proc. Geol Johnson Res. Found. 5th (1969, pub. 1971), 2, 195-204.

77. Oster O. J.Biol.Chem. (1975), 250, 7990-6.

78. Chiancone E. *et al*. J.Mol.Biol. (1972), 70, 675-88.

79. Hershberg R.D. PhD Thesis Univ. Pennyslavia 1973.

80. Novak R.F. *et al.* Mol Pharmacol. (1977) 13, 15-30.

81. Oldfield E. *et al.* J.Mol.Biol. (1975), 250, 6368-80.

82. Antonini E and Brunori M. "Haemoglobin and Myoglobin in their reactions with Ligands" (1971), American Elsevier, New York.

83. Frye J.S. and La Mar G.N. J.Am.Chem.Soc. (1975), 97, 3561-2.

84. Krishinamoorthi R. *et al.* J.Biol.Chem. (1984), 259, 8826-31.

85. La Mar G.N. *et al.* Nato ASI Ser A (1989) 183, 243-56.

86. Iizuka T. and Morishima I. Biochim.Biophys.Acta. (1974), 371, 1-13.

87. La Mar G.N. *et al.* Biochim.Biophys.Acta. (1980), 622, 210-18.

88. La Mar G.N. *et al.* J.Am.Chem.Soc. (1983), 105, 782-7.

89. La Mar G.N. in "Biological Applications of Magnetic Resonance" Ed. Shulman R.G. (1979), Academic Press New York.

90. Nelson M.J. *et al.* in "Haemoglobin and Oxygen Binding" Ed. Chien Ho (1982), Elsevier Holland.

91. Balch A.L. *et al.* J.Am.Chem.Soc. (1985), 107, 3006-7.

92. Bradbury J.A. and Carver J.A. J.C.S.Chem.Comm. (1981), 208-9.

93. Perutz M.F. Nature (1970), 228, 726-34.

94. Goff H. J.Am.Chem.Soc. (1977), 99, 7723-5.

95. Morishima I *et al.* Biochim.Biophys.Res.Commun. (1977), 78, 739-46.

96. Morishima I and Inubushi T. J.C.S.Chem.Comm. (1977), 77, 616-7.

97. Morishima I and Inubushi T. FEBS. Lett. (1977), 81, 57-60.

98. Behere D.V. *et al*. Biochim.Biophys.Acta. (1985), 862, 319-25.

99. Bradbury J.A. and Carver J.A. Biochem. (1984), 23, 4905-15.

100. Morishima I. and Hara Biochem. (1980), 19, 1569-75.

101. Morishima I. and Hara J.Am.Chem.Soc. (1982), 104, 6833-4.

102. Morishima I. and Hara Biochem. (1983), 22, 4102-7.

103. Jue T. and La Mar G.N. Biochim.Biophys.Res.Commun. (1984), 119, 640-5.

104. Levy M.J. *et al*. J.Biol.Chem. (1985), 260, 3694-8.

105. Constanidinis *et al*. Biochem. (1988), 27, 3069-76.

106. Takano T. J.Mol.Biol. (1977), 110, 557-68.

107. Jue T. *et al*. J.Am.Chem.Soc. (1983), 105,5701-3.

108. La Mar G.N. *et al*. J.Am.Chem.Soc. (1986), 108, 5568-73.

109. La Mar G.N. *et al*. J.Am.Chem.Soc. (1989), 111, 485-91.

110. La Mar G.N. *et al*. J.Am.Chem.Soc. (1984), 106, 6395-401.

111. Neya S. and Funasaki N. J.Biol.Chem. (1987), 262, 6725-8.

112. Neya S. and Funasaki N. Biochim.Biophys.Acta. (1988), 952, 150-7.

113. Springer B.A. *et al*. J.Biol.Chem. (1989), 264, 3057.

114. Gurd F.R.N. and Rothgeb T.M. Adv. Protein Chem. (1979), 33, 73-165.

115. Gurd F.R.N. *et al*. in "Biochemical Structure Determination by NMR", Ed. Bothner-By, Glickson and Sykes

(1982), Dekker, New York.

116. Maskalick D.G. PhD Thesis (1984), Univ. Microfilms Int. Ann Arbor Michigan.

117. Richmond T.J. and Richards F.M. J.Mol.Biol. (1978), 119, 537-55.

118. Cutnell J.D. *et al.* J.Am.Chem.Soc. (1981), 103, 3567-72.

119. Le-Comte J.T.J. and La Mar G.N. *Biochem. (1985), 24, 7388-95.*

120. Emerson S.D. et al. J.Am.Chem.Soc. (1988), 110, 4176-82.

121. Dayhoff M.O. "Atlas of Protein Sequence and Structure", (1969), Nat. Biomed. Res. Found, Silver Spring, Md.

122. Baltzer L. *et al.* J.C.S.Chem.Comm. (1985), 1040-1.

123. Lee H.C. *et al* J.Am.Chem.Soc. (1985), 107, 4087-8.

124. La Mar G.N. *et al.* Biochim.Biophys.Res.Commun. (1978), 82, 19-23.

125. La Mar G.N. *et al.*Biochim.Biophys.Res.Commun. (1985), 128, 628-33.

126. Toi H. *et al.* J.C.S.Chem.Comm. (1985), 1791-2.

127. Miller K.W. *et al* Proc. Natl. Acad. Sci. USA (1981), 78, 4946-9.

128. Tilton R.F. and Kuntz I.D. Biochem. (1982), 21, 6850-7.

129. Bondon *et al.* Biochim.Biophys.Acta. (1986), 872, 163-6.

130. Ikeda-Saito M. *et al.* J.Biol.Chem. (1978), 253, 7134-7.

131. Inubushi T. *et al.* Biochem. (1983), 22, 2904-7.

132. Kennedy M.A. and Ellis P.D. J.Am.Chem.Soc. (1989), 111, 3195-3208.

133. Lee H.C. and Oldfield E. J.Am.Chem.Soc. (1989), 111,

1584-90.

134. Yamamoto Y. FEBS. Lett. (1987), 222, 115-9.

135. Lecomte J.T.J. *et al*. Eur. Biophys. J. (1986), 13, 373-82.

136. Ramaprasad S. *et al*. J.Am.Chem.Soc. (1984), 106, 5330-5.

137. Sankar S. *et al*. Biochim.Biophys.Acta. (1987), 912, 220-9.

138. Wilbur D.J. and Allerhand A. FEBS. Lett. (1977), 79, 144-6

139. Keniry *et al* Biochem. (1983), 22, 1917-26.

140. Derome A.E. "Modern NMR Techniques for Chemistry Research" Pergamon Press (1987).

141. Redfield A.G. and Kunz S.D. J.Magn.Reson. (1975), 19, 250.

142. Nagayama K. and Wüthrich K. Eur.J.Biochem. (1981), 114, 365.

143. Bodenhausen G. and Freeman R. J.Magn.Reson. (1977), 28, 471.

144. Doddrell D.M. *et al*. J.Magn.Reson. (1982), 48, 323-7.

145. Bax A. and Freeman R. J.Magn.Reson. (1981, 44, 542.

146. Morris G.A. and Freeman R. J.Am.Chem.Soc. (1979), 101, 760-2.

147. La Mar G.N. and Walker F.A. J.Am.Chem.Soc. (1973), 95, 1782-90.

148. Traylor T.G. and Berzinis A.P. J.Am.Chem.Soc. (1980), 102, 2844-6.

149. Dalvit C. and Wright P.E. J.Mol.Biol. (1987), 194, 313-

27.

150. Kuriyan J. *et al.* J.Mol.Biol. (1986), 192, 133-54.

151. The same assignments were independently obtained by use of one and two dimensional NOE techniques by La Mar G.N. *et al.*[85].

152. Emerson S.D. PhD Thesis Univ. of California (Davis) 1988.

153. La Mar, G.N.Harrocks, W. DeW and R.H.Holm NMR of paramagnetic molecules : Principles and applications Academic New York 1973

154. R.A.Dwek, Nuclear Magnetic resonance in biochemistry (applications to enzyme systems) Clarendon Press, Oxford, 1973

155. Johnson C.E. and Bovey F.A. J.Chem.Phys. (1958),29, 1012.

156. Shulman R.G., Ogawa S. Wüthrich K. and Blumberg W. Science (1969), 165, 251.

157. Wüthrich K. *et al.* Proc. Nat. Acad. Sci. US (1968), 60, 373.

158. Williamson M.P., Madison V.S. Biochem. (1990),29,2895.

159. Roy M. *et al.* J.Biol.Chem. (1990), 265, 5448-52.

160. Fesik S.W. *et al.* J.Am.Chem.Soc. (1989), 111, 770.

161. Wijmenga S.S. *et al.* J.Magn.Reson. (1989), 84, 684.

162. Bax A. *et al.* J.Magn.Reson. (1990), 88, 425-31.

163. Kay L.E. J.Am.Chem.Soc. (1990), 112, 888

164. Clore G.M. *et al.* Biochem. (1991), 30, 2815-28.

165. Ward A.J. and Nelson S.J. Biophys J. (1991), 59, 1101-2.

166. Nelson S.J. *et al.* Biophys J. (1991), 59, 1113-22.

167. Gippert G.P. *et al.* Biochem Pharm. (1990), 40, 15-22.

168. Williamson M.P Biopolymers (1990), 29, 1423-31.

169. Osapay K. and Case D.A. J.Am.Chem.Soc. (1991), 113, 9436-44.

170. Pastore A. and Sandek V. J.Magn.Reson. (1990), 90, 165-76.

171. Arrowsmith C.H. *et al.* Makromol. Chem. Macromol. Symp. (1990), 34, 33-46.

172. Le Master D.M. Q. Rev. Biophys. (1990), 23, 133-74.

173. Derrick J.P. *et al.* Febs. Lett. (1991), 280, 125-8.

174. Bogusky M.J. *et al.* J.Magn.Reson. (1990), 86, 11-29.

175. Bogusky M.J. *et al.* Biochim.Biophys.Res.Commun. (1985), 127, 540.

176. Fesik F.W. *et al.* Biochem. (1988), 27, 8297.

177. Oldfield E. *et al.* Biochem. (1991), 30, 2333-47.

178. Oldfield E. *et al.* J.Am.Chem.Soc. (1991), 113, 8680-5.

179. Yamamoto Y. and Chûjô R. J.C.S.Chem.Comm. (1992), 87-9.

180. Lecomte J.T.J. and Cocco M.J. Biochem. (1990), 29, 11057-67.

181. Cocco M.J. and Lecomte J.T.J. Biochem. (1990), 29, 11067-72.

182. Deeb S.R. and Peyton D.H. J.Biol.Chem. (1991), 266, 3728-33.

183. Deeb S.R. and Peyton D.H. Biochem. (1992), 31, 468-74.

184. Hauksson J.B. *et al.* J.Am.Chem.Soc. (1990), 112, 8315-23.

185. La Mar G.N. *et al.* J.Am.Chem.Soc. (1991), 113, 1544-50.

186. Simonneaux *et al.* Biochim.Biophys.Acta. (1989), 999, 42-5.

187. Adachi S-i *et al.* Biochim.Biophys.Res.Commun. (1991),

180, 138-44.

188. Yamamoto Y. *et al*. J.C.S.Chem.Comm. (1990), 1556-7.

189. Yamamoto Y. *et al*. Febs. Lett. (1990), 264, 113-6.

190. Yamamoto Y. *et al*. Biochim.Biophys.Acta. (1992), 1120, 173-82.

191. Sultis S.M. and Strouse C.E. J.Am.Chem.Soc. (1988), 110, 2824-9.

192. Emerson S.D. and La Mar G.N. Biochem. (1990), 29, 1545-56.

193. Emerson S.D. and La Mar G.N. Biochem. (1990), 29, 1556-65.

194. Yu L.P. *et al*. J.Am.Chem.Soc. (1990), 112, 9527-34.

195. Reyton D.H. Biochim.Biophys.Res.Commun. (1991), 175,5159.