Nuclear Overhauser effects and the assignment of the proton NMR spectra of proteins

Christopher M. Dobson, Martin A. Howarth and Christina Redfield

Inorganic Chemistry Laboratory, South Parks Road, Oxford OX1 3QR, England

Received 7 September 1984

A systematic procedure is outlined to reveal assignments in the ^{1}H NMR spectrum of a protein that are consistent with a limited set of nuclear Overhauser effect data and with specific assumptions. The results of applying this to a group of resonances in the spectrum of hen lysozyme, most of which were previously assigned by independent methods, are described. The number of possible assignment schemes obtained depends on the distance constraints chosen for the Overhauser effect data, and the investigation of this dependence is proposed as a method of assessing the confidence that can be placed in individual assignments obtained in this way.

^{1}H NMR Lysozyme NOE Assignment

1. INTRODUCTION

The importance of nuclear Overhauser enhancement (NOE) measurement in the process of assigning the ^{1}H NMR spectra of proteins is now well established [1–4]. The existence of an Overhauser effect between a pair of protons, detected by one or two-dimensional techniques, establishes them as close together in the protein structure [5,6]. In the simplest case, where initial rate measurements are possible and where the spin system can be described as undergoing simple isotropic motion, the relative magnitudes of NOE effects depend on $1/r^6$, where r is the distance separating the two protons [1,7]. In proteins, however, there are deviations from this simple picture. It is in general not possible to measure detailed time dependences of NOE effects for any except the closest protons because of the large number of spins in the system and because the signal-to-noise ratio is rarely adequate. In addition, the existence of internal motions causes variations in effective correlation times for different pairs of protons in the molecule [6–8]. Nevertheless, theoretical analysis and experimental results provide strong support for a high correlation of the magnitude of the NOE effect with interproton distance in those cases where detailed studies have been carried out [6].

Several approaches have been put forward to obtain assignment information from NOE measurements. In one, use is made of the fact that the amide, H^{α}, and H^{β} protons of a given residue $(i-1)$ are always close to the amide proton of the adjacent residue (i) in the protein sequence, whatever the value of the torsion angles ψ and ϕ [4,9]. Thus, with a sufficiently detailed knowledge of the spin systems present, it is possible to derive assignments from NOE measurements, moving along the protein in a sequential manner. This approach has been developed and demonstrated for peptides and small proteins [10–12].

For many proteins, however, it is not yet feasible to obtain as complete a description of the NMR spectrum as is needed for this procedure. Further, it may not be necessary if only a localized region of the protein, for example the active site of an enzyme, is of interest. An alternative approach to

Abbreviations: NOE, nuclear Overhauser enhancement; lysozyme, hen egg-white lysozyme (EC 3.2.1.17)

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

assignment which has been used in studies of these molecules is to assume that the average structure of the protein in solution is essentially the same as that found in the crystalline state [1,13]. Inter-proton distances can, therefore, be calculated from the crystal structure and used to interpret NOE data in terms of resonance assignments [1,14–16]. In this approach the degree of similarity of the solution and crystal structures can subsequently be examined [17].

Here, we consider a system of resonances of lysozyme, the majority of which have been assigned previously by methods not involving NOE effects to protons in a region of the molecule called the hydrophobic box [1]. An extensive set of NOE measurements exists for these resonances, and these are used in conjunction with the high-resolution crystal structure of the protein to investigate some aspects of this second approach to assignment. In particular, a computer-based search procedure is employed to explore the effect of varying the criteria used for defining possible assignments.

2. METHODS

Two types of information are used in the assignment analysis. First, spin-spin coupling is used to identify groups of resonances from individual residues. Consideration of the coupling patterns and chemical shifts of these resonances permits some restrictions to be placed on the type of residue to which each group of coupled resonances belongs. Second, NOE data are used to provide information about the proximity in the molecular structure of the residues from which different sets of resonances arise. In conjunction with information about the structure, this provides further restrictions on the acceptable assignments of each group of resonances.

The degree to which the first stage in this procedure restricts possible assignments depends upon the completeness of the coupling information. In addition, it is necessary to consider that more than one set of coupled resonances may arise from the same residue, and that in some cases it will not be possible to correlate these sets. This will occur, for example, if the residue contains a heavy atom without a directly bonded proton; 4-bond coupling is usually too small to resolve in protein spectra. Such residues include methionine and all the aromatic residues. It can also occur simply when 3-bond coupling cannot be resolved, for example because the coupling constant is small or a resonance is broad.

The degree to which the second stage in this procedure restricts possible assignment schemes depends on the number of NOE effects observed, and on the constraints that can be placed upon the proximity of protons between which an NOE effect is observed. The analysis adopted here involves a computer program that operates as follows. Each NOE effect is considered in turn, and all assignment pairs consistent with the restrictions on residue type derived from the coupling and chemical shift data are identified. A search is then carried out through the protein structure to determine which of these pairs of protons are within a specified distance of each other. This limiting distance is a parameter which may be varied. In the final stage of the analysis the lists of possible assignment pairs for each NOE experiment are used to find assignment schemes for the whole group of resonances which are self-consistent. A valid assignment for a given set of resonances must appear in the list of assignment pairs for every NOE effect involving resonances of that set.

The data used in the analysis discussed in this paper are shown in fig.1. The coupling and NOE information has been described previously [1]. In two cases, labelled A and B, the sets of resonances are of aliphatic protons, and could arise from the same residue as a number of the other sets of resonances for the reasons outlined above. This possibility was permitted, but not required, in the computer program used here. Proton coordinates based on the refined tetragonal X-ray structure of lysozyme ([18]; Grace, D.E.P., Phillips, D.C., Artymiuk, P.J., unpublished), were obtained using a standard procedure [19]. The distances involving methyl groups and aromatic rings of tyrosine and phenylalanine were $1/r^6$ averages to take into account the rotational behaviour observed for these residues. In addition, average distances were used for methylene protons because of the difficulty of distinguishing the protons within these pairs. The computer analysis was carried out on the VAX-11/780 computer of the Oxford University Computing Service.

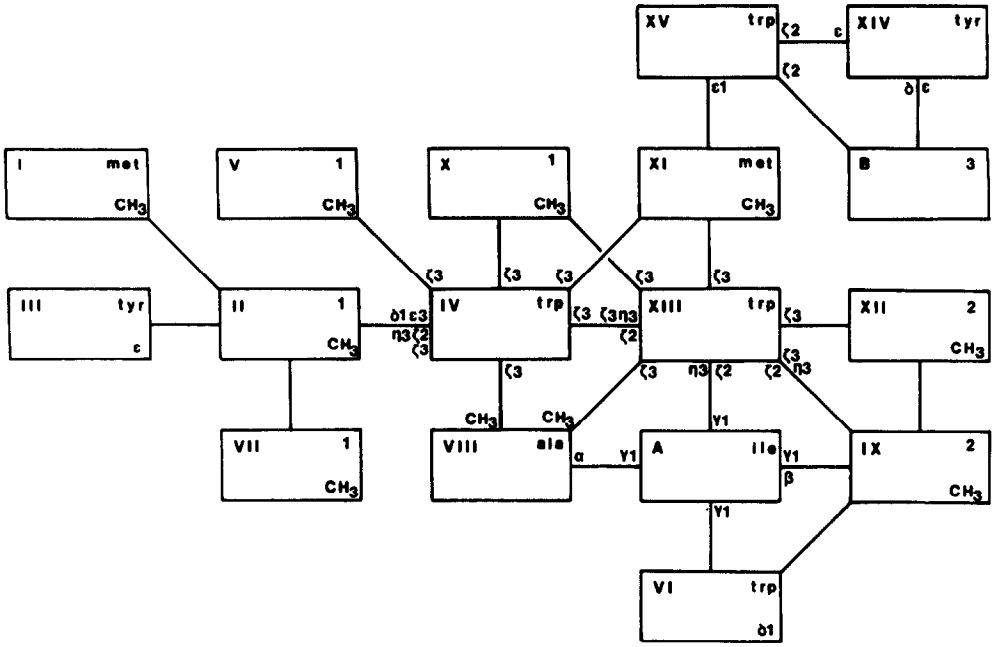

Fig.1. NOE effects used in the analysis of the system of residues studied in this work. Each box represents a resonance or a group of resonances which have been shown to belong to the same spin system. The lines joining the boxes represent observed NOE effects, the letters indicating which protons of the coupled set are involved. In cases where all NOE effects to a spin system involve the same resonance, the proton type is given in the lower right-hand corner of the box. Conclusions about residue type from spin-spin coupling patterns for each set of coupled resonances are indicated in the upper right-hand corner of each box. Type 1 indicates that the spin system may be assigned to any leucine or valine residue, type 2 to any isoleucine, alanine or threonine residue, and type 3 indicates that the resonance may correspond to any methine or methylene proton. The coupling patterns of all other spin systems are characteristic of a single residue type, as indicated.

3. RESULTS AND DISCUSSION

The system of NOE effects shown in fig.1 was analysed as described in section 2. The distance constraints were varied in steps from 0 Å up to 8 Å, and the acceptable assignment schemes obtained for the various values of the distance constraints are shown in table 1. The number of acceptable schemes is plotted as a function of the distance constraints in fig.2. No assignment schemes were found to be consistent with the whole set of data unless the constraints were 5.9 Å or larger. At this value, one assignment scheme is obtained. As the distance constraints were relaxed further, no other acceptable assignment schemes were found until 6.3 Å was reached. The new allowed scheme differs from the previous one at 5.9 Å in the assignment of just one residue, VI. An increasing number of schemes becomes possible as the constraints are relaxed to 8 Å. At larger values than this, the requirements for numerical handling of the computing system become very large. At 8 Å, there are a total of 24 assignment schemes consistent with the coupling and NOE data but the assignments of only 6 of the 17 residues differ between these schemes. Hence, although the system as a whole has many acceptable assignment schemes, for 11 residues an alternative assignment is not found.

If the crystal and solution structures are identical then at least one NOE effect would have to arise between protons separated by a distance of greater than 8 Å in order for one of these 11 residues to have a different assignment to that given in table 1. Alternatively, if the crystal and solution structures differ, and if 6 Å were taken as a limiting distance beyond which NOE effects are not observed, then differences between the two structures would have to exceed 2 Å. One of these possibili-

Table 1

Assignments obtained for the residues of fig.1 using specified distance constraints

Distance constraint (Å)	Residue																
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	A	B
	Assignments[a]																
5.9	Met 12	Leu 17	Tyr 20	Trp 28	Leu 56	Trp 63	Val 92	Ala 95	Ile 98	Val 99	Met 105	Ala 107	Trp 108	Tyr 23	Trp 111	Ile 98	Met 105
6.3						Trp 62											
6.5					Val 99					Leu 56							
7.0																Ile 58	Asn 27
7.5[b]					Val 92		Leu 56 Val 99										

[a] Assignments are given for all the residues at a distance constraint of 5.9 Å. Alternative assignments found to be acceptable as the distance constraints are increased, are then listed subsequently

[b] No additional acceptable assignments were found with a constraint of 8.0 Å

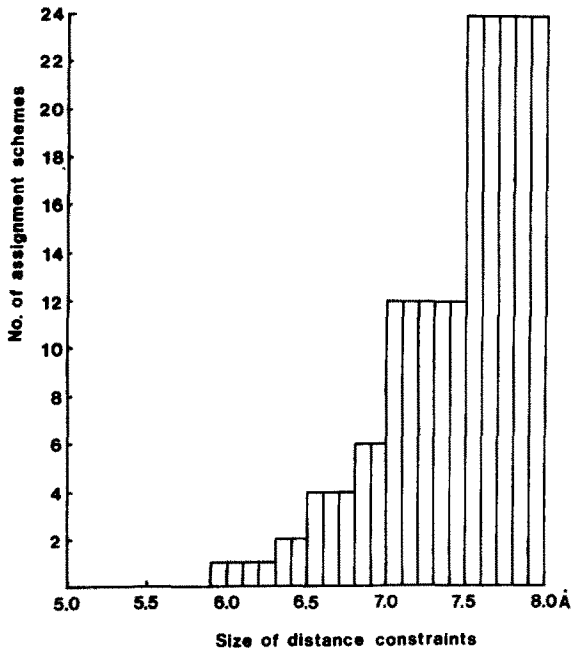

Fig.2. Histogram of the number assignment schemes consistent with the distance constraints as the constraints are varied up to 8 Å.

ties is also implied if any of the other 6 residues (V, VI, VII, X, A and B) are not to be assigned to one of the alternatives given in table 1. It is interesting to note that the assignment scheme which satisfies the distance constraints most easily (5.9 Å) agrees in every detail with the set of assignments made previously which was not based on consideration of NOE data [20–22].

The results of this analysis suggest that the investigation of the effect of varying constraints on the possible assignments for a set of residues can provide information about the level of confidence that may be held in assignments made using a given set of experimental data. The definition of a level of confidence in particular assignments is likely to be useful as it indicates when further experimental information is needed to support specific assignments. A general decrease in the size of the NOE effects with increasing interproton distance is the behaviour anticipated theoretically and found experimentally in [5,6]. For every residue, two distances may be defined on the basis of the analysis given. First, the distance to which the constraints must be relaxed in order to obtain an acceptable

assignment for that residue, and second the shortest distance at which an alternative assignment becomes acceptable. The level of confidence which may be held in an assignment made on this basis will increase as the difference between these two distances increases. The experimental data in table 1, therefore, define residues V, VI, VII, X, A and B less completely than the others in terms of assignment possibilities discussed here. One may, however, be confident that the assignments of these residues will be among the alternatives shown in table 1. Twenty-four possible assignment schemes at 8 Å would seem to imply a considerable lack of definition of the residues, but compared to the number of possible schemes consistent with the coupling data alone (about 4×10^{14}), this is an extremely small number.

A histogram of the interproton distances found for the assignment scheme consistent with the shortest distance constraints is shown in fig.3. Almost all the NOE effects involve side chain rather than main chain protons which are separated by distances generally much larger than those involved in the sequential assignment method. Indeed, in the sequential assignment method NOE effects are generally assumed to be from protons separated by <3.5 Å [9] whilst all the residues assigned here have at least one NOE involving a distance of >3.5 Å. Nevertheless, the unique assignment scheme obtained with the tightest distance constraints is precisely that proposed on the basis of independent chemically based procedures.

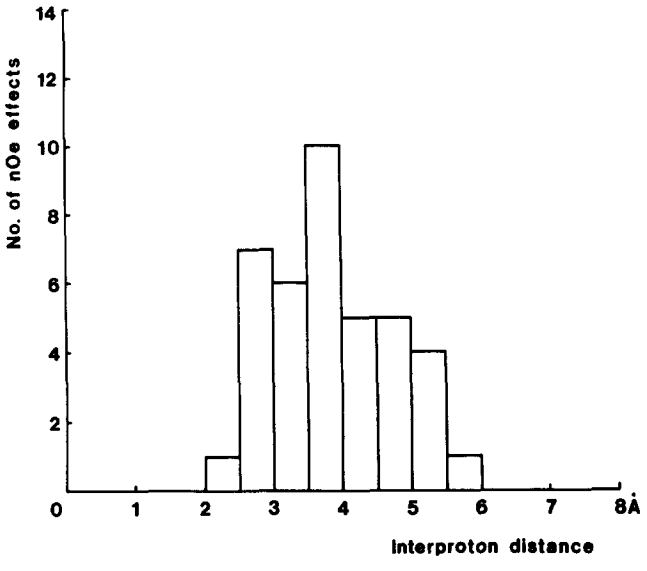

Fig.3. Histogram of the interproton distances found for the 39 NOE effects analysed in the system shown in fig.1.

The approach described in this paper can readily be adapted to the needs of specific assignment problems. In particular, in the coupling data analysis, the set of allowed residue types which a spin system may be assigned to can be restricted to any combination of the amino acid residues found in the protein. Such a detailed specification of the restrictions derived from the coupling data has been found to be crucial in limiting the numbers of acceptable assignment schemes and it also leads to complete flexibility at this stage of the assignment process. In the NOE data analysis, information about the magnitudes of NOE effects can be incorporated for example by specifying different distance contraints for large and small NOE effects. Recently this systematic approach to the analysis of NOE data has been used in the assignment of a substantial number of amide, H^{α} and H^{β} resonances in the spectrum of hen lysozyme [17].

ACKNOWLEDGEMENTS

Ths work is supported by the Science and Engineering Research Council. C.R. acknowledges a Fellowship from Bell Telephone Laboratories. We thank C. Bethel, M. Delepierre and F.M. Poulsen for valuable discussions. C.M.D. is a member of the Oxford Enzyme Group.

REFERENCES

[1] Poulsen, F.M., Hoch, J.C. and Dobson, C.M. (1980) Biochemistry 19, 2597–2607.
[2] Zuiderweg, E.R.P., Kaptein, R. and Wüthrich, K. (1983) Eur. J. Biochem. 137, 279–292.
[3] Inagaki, F., Tamiya, N., Miyazawa, T. and Williams, R.J.P. (1981) Eur. J. Biochem. 118, 621–625.
[4] Wagner, G. and Wüthrich, K. (1982) J. Mol. Biol. 155, 347–366.
[5] Noggle, J.H. and Schirmer, R.E. (1971) The Nuclear Overhauser Effect, Academic Press, New York.

[6] Delepierre, M., Dobson, C.M., Hoch, J.C., Olejniczak, E.T., Poulsen, F.M., Ratcliffe, R.G. and Redfield, C. (1981) in: Biomolecular Stereodynamics (Sarma, R.H. ed.) Adeneine Press, New York.
[7] Wagner, G. and Wüthrich, K. (1979) J. Magn. Reson. 33, 675–680.
[8] Olejniczak, E.T., Poulsen, F.M. and Dobson, C.M. (1981) J. Am. Chem. Soc. 103, 6574–6580.
[9] Billeter, M., Braun, W. and Wüthrich, K. (1982) J. Mol. Biol. 155, 321–346.
[10] Dubs, A., Wagner, G. and Wüthrich, K. (1979) Biochim. Biophys. Acta 577, 177–194.
[11] Inagaki, F., Clayden, N.J., Tamiya, N. and Willams, R.J.P. (1982) Eur. J. Biochem. 123, 99–104.
[12] Nagayama, K. (1981) Adv. Biophys. 14, 139–204.
[13] Dobson, C.M., Hoch, J.C., Olejniczak, E.T. and Poulsen, F.M. (1980) Biophys. J. 32, 625–628.
[14] Canters, G.W., Hill, H.A.O., Kitchen, N.A. and Adman, T. (1984) Eur. J. Biochem. 138, 141–152.
[15] Delepierre, M., Dobson, C.M. and Poulsen, F.M. (1982) Biochemistry 21, 4756–4761.
[16] Redfield, C., Poulsen, F.M. and Dobson, C.M. (1982) Eur. J. Biochem. 128, 527–531.
[17] Delepierre, M., Dobson, C.M., Howarth, M.A. and Poulsen, F.M. (1984) Eur. J. Biochem., in press.
[18] Grace, D.E.P. (1980) D. Phil. Thesis, Oxford University.
[19] Hoch, J.C. (1983) D. Phil. Thesis, Harvard University.
[20] Campbell, I.D., Dobson, C.M. and Williams, R.J.P. (1975) Proc. R. Soc. Lond. A 345, 41–59.
[21] Dobson, C.M., Ferguson, S.J., Poulsen, F.M. and Williams, R.J.P. (1978) Eur. J. Biochem. 92, 99–103.
[22] Cassels, R., Dobson, C.M., Poulsen, F.M. and Williams, R.J.P. (1978) Eur. J. Biochem. 92, 81–97.