Chemical shifts of aromatic protons in protein NMR spectra

Christina Redfield, Jeffrey C. Hoch* and Christopher M. Dobson+

+*Inorganic Chemistry Laboratory, South Parks Road, Oxford OX1 3QR, England* and *Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA*

Received 13 June 1983

Experimental secondary shifts have been compared with calculated ring current shifts for the aromatic protons of lysozyme. The root mean square differences were calculated for all possible permutations of the assignment scheme. The experimentally determined assignment scheme was found to give the best agreement with the ring current shift calculations. Similar results were obtained in the case of bovine pancreatic trypsin inhibitor and cytochrome *c*. The use of the degree of correlation between experimental and predicted NMR parameters as a basis for making assignments is examined.

Aromatic proton resonance *Lysozyme* *NMR spectra* *Secondary shift* *Ring current shift* *Chemical shift assignment*

1. INTRODUCTION

The chemical shift values of protons in a globular protein differ from those in an unstructured polypeptide primarily because of contributions from through-space interactions between a proton and the residues in its vicinity [1]. The secondary shifts caused by these interactions are thus extremely sensitive to, and characteristic of, the conformation of the protein. In diamagnetic proteins the shifts are generally less than about 1 ppm, but may be as large as 5 ppm, and 3 mechanisms are thought to be important. Shifts can arise from the anisotropy of the magnetic susceptibility associated with ring currents in aromatic groups, from the local anisotropy associated with other groups (e.g., carbonyls) or atoms, and from the perturbations to electron distributions arising from the electric fields associated with charged groups [2,3].

+ To whom correspondence should be addressed

* Present address: Rowland Institute for Science, 100 Cambridge Parkway, Cambridge, MA 02142, USA

Abbreviations: ppm, parts per million; BPTI, bovine pancreatic trypsin inhibitor; NMR, nuclear magnetic resonance; RMS, root mean square

Theoretical models for calculating the magnitude of ring current shifts for a given molecular conformation are well developed [2]. Provided that resonance assignments are known, these predictions can be compared with experimental values of secondary shifts. From the comparison made so far, chiefly with lysozyme, bovine pancreatic trypsin inhibitor (BPTI) and cytochrome *c*, there is a reasonable correlation between the observed and calculated values for protons bonded to carbon [4–6]. The correlation appears to be particularly satisfactory for methyl group resonances. For aromatic proton resonances the situation is less clear; ring current calculations have not, for example, provided any explanation for the large shifts of the aromatic protons of BPTI [5]. Other contributions are presumably important here, but methods of calculation are not sufficiently advanced for these to be predicted satisfactorily.

Recently, we completed the assignment of all aromatic proton resonances in the NMR spectrum of lysozyme [7]. With this information, a more detailed examination of the chemical shifts of aromatic protons in proteins is possible. This paper reports the results of such a study, which give insight into the value of ring current shift calculations.

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

2. METHODS

Ring current shift values for the aromatic residues of lysozyme and BPTI were calculated using the appropriate crystallographic coordinates and a FORTRAN ring current shift program [3]. In the case of lysozyme, ring current shift values were calculated for two coordinate sets, the refined tetragonal crystal structure [8] and the triclinic crystal structure [9]. Ring current shift calculations were carried out using 3 ring current shift models, the classical dipolar equation [10], the semi-classical Johnson–Bovey equation in [11] and the quantum mechanical Haigh–Mallion equation in [12]. The calculations were carried out using the ring current intensity factors proposed in [13] and those more recently suggested in [4]. Ring current shift values for tuna cytochrome *c* were obtained from [6].

The experimental chemical shift values for lysozyme, BPTI and tuna cytochrome *c* used in this study were obtained from [7,14–16]. Secondary shifts were calculated as the difference between the experimental chemical shift values and basis shift values for the amino acid in a random coil polypeptide. The random coil basis shift values for tyrosine and tryptophan in [17] and the values for phenylalanine in [6] were used in these calculations.

The analysis used here is carried out as follows. For each amino acid type (tyrosine, tryptophan and phenylalanine) it is assumed that all the spin systems have been identified and that each resonance has been assigned to a specific proton type (e.g., H^{δ}, H^{ϵ} or H^{ζ} for phenylalanine). A FORTRAN program is used to compare the theoretical ring current shift values with the experimental secondary shift values, and the root mean square (RMS) difference is calculated for each of the *n!* assignment schemes possible for *n* spin systems of a given amino acid type. The assignment schemes are then ordered in terms of the quality of fit of the experimental data with the theoretical predictions as measured by the RMS difference. The RMS difference is plotted for each of the *n!* possible assignment schemes; the best fit assignment scheme is given rank 1 and the worst assignment scheme is given rank *n!* on the *x*-axis. It should be noted that in the case of slowly rotating tyrosine and phenylalanine residues, or when the specific proton type cannot be assigned unambiguously (e.g., H^{δ} or H^{ϵ} for tyrosine), the number of possible assignment schemes may be greater than *n!*.

3. RESULTS

The RMS differences between observed secondary shifts and calculated ring current shifts for the aromatic protons of lysozyme are listed in table 1. Similar values for BPTI and cytochrome *c* are evident from fig.3 and 4. It can be seen that the RMS difference varies little with the method used for the ring current shift calculations; in this paper the Johnson–Bovey equation has been used [2,3]. The use of empirical ring current intensity factors for tryptophan in [4] leads to improved agreement for the tryptophan residues but to worse agreement for the tyrosine and phenylalanine residues of lysozyme. The overall RMS difference of nearly 0.2 ppm reported here and evidence from other studies indicates that other contributions to the secondary shift such as the local anisotropy associated with groups such as carbonyls and electrostatic effects may be important [2–6]. Theoretical models for these other contributions are not as well developed as those for ring current shifts, and the success of calculations incorporating these contributions in improving agreement with experiment has been limited [2,3].

All 59 aromatic protons of lysozyme have been assigned without reference to ring current shift

Table 1

Root mean square difference between calculated and observed secondary shifts in ppm for the aromatic residues of lysozyme

	Johnson-Bovey[a]	Haigh-Mallion[a]	Dipole[a]	Johnson-Bovey[b]
Tyrosine	0.08	0.08	0.06	0.08
Tryptophan	0.22	0.22	0.21	0.18
Phenylalanine	0.18	0.18	0.18	0.21
Overall	0.19	0.18	0.18	0.17

[a] Calculation carried out using ring current intensity factors, as in [13]

[b] Calculation carried out using empirical tryptophan ring current intensity factor, as in [4]

calculations [7]. As illustrated above, these assignments give an RMS difference of 0.18 ppm with the calculated ring current shifts. In order to assess the quality of fit of the experimental data with the calculated shifts, the RMS difference was calculated for all possible permutations of the assignment scheme (*n!* possible assignment schemes for *n* spin systems). The results of this calculation are illustrated in fig.1 for the refined tetragonal coordinates of lysozyme. It can be seen that in all cases the experimentally determined assignment scheme gives the best agreement (lowest RMS difference) with the calculated ring current shifts. The same type of analysis was also applied to the lysozyme triclinic coordinate set. In the case of the tryptophan and tyrosine residues the experimental assignment scheme again gives the lowest RMS difference. In the case of phenylalanine residues, however, the experimental assignments give the worst agreement as indicated in fig.2. The 3 phenylalanine residues of lysozyme are located in the same region of the structure in close proximity to a tryptophan residue. The ring current shift values for these phenylalanine residues are determined by the relative orientations of these 4 aromatic rings. These 4 aromatic groups

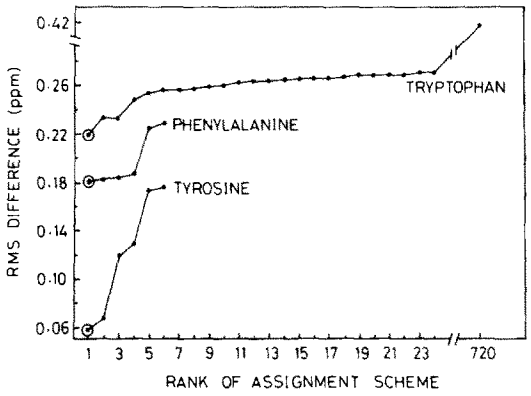

Fig.1. Plot of the RMS difference between experimental secondary shifts and calculated ring current shifts for each of the *n!* assignment schemes for the tryptophan, phenylalanine and tyrosine residues of lysozyme. The possible assignment schemes are ranked in terms of the quality of fit between experimental secondary shifts and calculated ring current shifts as measured by the root mean square difference. In this and subsequent figures the experimentally determined assignment schemes are circled.

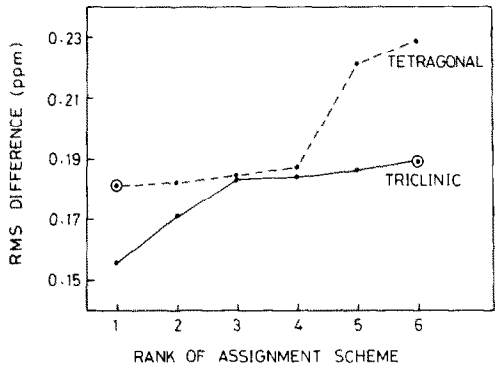

Fig.2. Plot of the RMS difference between experimental secondary shifts and calculated ring current shifts for the 6 possible assignment schemes for the 3 phenylalanine residues in lysozyme. Ring current shifts were calculated using the refined tetragonal crystal structure [8] and the refined triclinic crystal structure [9].

do not contribute significantly to the ring current values of any of the other aromatic protons in lysozyme. The result illustrated in fig.2 shows that differences in the two crystal structures in the region of the 3 phenylalanine residues lead to significant differences in the ring current shift calculations.

Complete assignment of the aromatic resonances of BPTI and of the phenylalanine resonances of tuna ferrocytochrome *c* have been made independently of ring current shift calculations [14–16]. A similar ring current analysis was carried out on these proteins in order to determine whether the results presented above are a general phenomenon or are characteristic of lysozyme alone. The results obtained for BPTI and cytochrome *c* are illustrated in fig.3 and 4. In the case of the 4 phenylalanine residues of BPTI, the experimentally determined assignment scheme gives the best agreement with ring current shift calculations, although the RMS difference of 0.26 ppm is greater than that obtained for lysozyme. In the case of the 4 tyrosine residues, the experimentally obtained assignment scheme is ranked 11th out of the 24 possible assignment schemes. It can be seen, however, for tyrosine that the RMS difference varies little with the assignment scheme.

In the case of the 3 phenylalanine residues of tuna cytochrome *c*, the experimentally determined assignment scheme gives the best agreement with

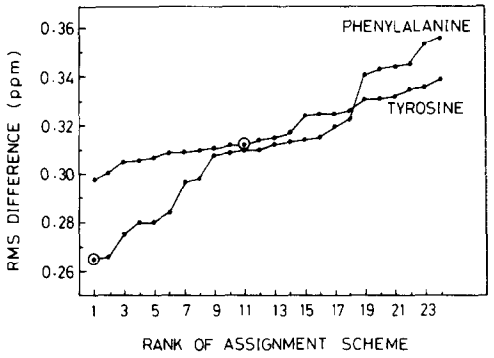

Fig.3. Plot of the RMS difference between experimental secondary shifts and calculated ring current shifts for each of the *n!* assignment schemes for the phenylalanine and tyrosine residues of BPTI.

the ring current shift calculations. A larger number of resonance assignments have been made for horse ferrocytochrome *c* but as yet a high resolution crystal structure of this protein is not available. Horse cytochrome *c* NMR data have been compared with ring current shift data calculated from the tuna crystal structure in [6]. In the light of differences in resonance chemical shift values between the two proteins and the occurrence of a large number of amino acid substitutions [15,16], some involving aromatic residues, a more quantitative comparison of horse cytochrome *c* secondary shift values and tuna cytochrome *c* ring current shift calculations here does not seem justified.

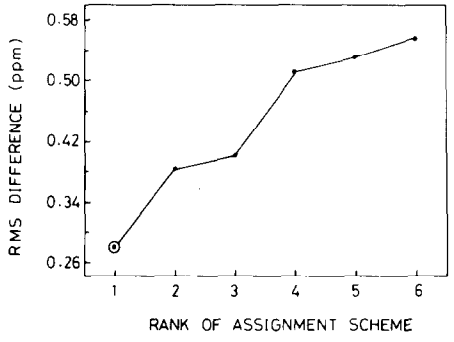

Fig.4. Plot of the RMS difference between experimental secondary shifts and calculated ring current shifts for each of the 6 assignment schemes for the 3 phenylalanine residues of tuna ferrocytochrome *c*.

4. DISCUSSION

Despite the rather large RMS differences between observed secondary shifts and calculated ring current shifts for the aromatic proton resonances of the 3 proteins considered here, there are several interesting conclusions which can be drawn from our results. First, in no case is the RMS difference for an incorrect assignment scheme significantly better than for the correct scheme. This strongly implies that the ring current shifts are indeed a major contribution to the secondary shifts of the aromatic protons, and that the methods of calculation are reasonable. Only in the case of the tyrosine residues of BPTI is it apparent that other contributions to the shifts are dominant. That this is the case is, however, amply revealed by the insensitivity of the large RMS difference to the permutation of possible assignments. It has been suggested that the contributions of local anisotropy effects, chiefly from the carbonyl groups, is the explanation of this [2,5].

The results of this study do not justify the use of ring current shift calculations by themselves as a means of making firm assignments. Although the results show that the correct assignment scheme gives the lowest RMS difference, this is not a sufficient criterion for making assignments since the RMS difference for many different assignment schemes are rather similar. The RMS difference of about 0.2 ppm obtained in this study is of the order of the average secondary shift observed for the aromatic protons of the 3 proteins studied. In isolated cases, such as that of Trp 28 in lysozyme, where secondary shifts are large, of the order of 0.8 ppm, assignments based on ring current shift calculations can perhaps be made with confidence. However, in the majority of cases where secondary shifts are small, of the order of the RMS difference, assignments cannot reliably be based on ring current shift calculations. Other methods, notably those involving nuclear Overhauser effects, are needed here [18]. The approach illustrated here of examining quantitatively the correlation between experimental and predicted NMR parameters is, however, likely to be of general applicability. If the degree of correlation is used as a basis for making assignments it is essential to assess the quality of the correlation for the proposed assignment relative to other possible assignment schemes.

Our results do suggest that the correct assignment scheme would be expected to predict the secondary shifts well compared with most incorrect schemes. If the RMS difference is significantly greater for the proposed assignment scheme than for other schemes, there are grounds for close examination of the reasons for this. Apart from an error in the assignment process, it could indicate a significant and unusual difference between the solution conformation and the coordinate set used as the basis for the calculations. Such a situation provides an explanation for the observations with the phenylalanine residues of lysozyme. The solution conformation in the region of the phenylalanine residues appears to be closer to that observed in the tetragonal crystals than in the triclinic crystals.

ACKNOWLEDGEMENTS

This work is supported by the Science and Engineering Research Council. This is a contribution from the Oxford Enzyme Group. C.R. acknowledges support from Bell Telephone Laboratories and from Wellesley College. We thank Professor L.H. Jensen for providing the triclinic lysozyme coordinates and Professor D.C. Phillips for providing the tetragonal lysozyme coordinates. We thank Professor M. Karplus for valuable discussions, and Dr G.R. Moore for providing the cytochrome *c* assignment data prior to publication.

REFERENCES

[1] Campbell, I.D. and Dobson, C.M. (1979) in: Methods of Biochemical Analysis (Glick, D. ed) vol.25, pp.1–133, Wiley, New York.
[2] Perkins, S.J. (1982) in: Biological Magnetic Resonance (Berliner, L.J. and Reuben, J. eds) vol.4, pp.193–336, Plenum, New York.
[3] Hoch, J.C. (1983) PhD Thesis, Harvard University.
[4] Perkins, S.J. and Dwek, R.A. (1980) Biochemistry 19, 245–258.
[5] Perkins, S.J. and Wüthrich, K. (1979) Biochim. Biophys. Acta 576, 409–423.
[6] Perkins, S.J. (1980) J. Magn. Reson. 38, 297–312.
[7] Redfield, C., Poulsen, F.M. and Dobson, C.M. (1982) Eur. J. Biochem. 128, 527–531.
[8] Grace, D.E.P., Phillips, D.C. and Artymiuk, P.J. (1983) unpublished.
[9] Hodson, J.M., Sieker, L.C. and Jensen, L.H. (1975) Am. Cryst. Assn. abst.3, 16.
[10] Pople, J.A. (1956) J. Chem. Phys. 24, 1111.
[11] Johnson, C.E. and Bovey, F.A. (1958) J. Chem. Phys. 29, 1012–1014.
[12] Haigh, C.W. and Mallion, R.B. (1971) Mol. Phys. 22, 955–970.
[13] Giessner-Prettre, C. and Pullman, B. (1969) CR Hebd. Seances Acad. Sci. Ser. D 268, 1115–1117.
[14] Wüthrich, K. and Wagner, G. (1979) J. Mol. Biol. 130, 1–18.
[15] Moore, G.R. and Williams, R.J.P. (1980) Eur. J. Biochem. 103, 493–502.
[16] Moore, G.R. and Williams, G. (1983) unpublished.
[17] Bundi, A. and Wüthrich, K. (1979) Biopolymers 18, 825.
[18] 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) pp.237–253, Academic Press, New York.