Volume 302, number 2, 185–188 FEBS 11016 May 1992

The application of ^{1}H NMR chemical shift calculations to diastereotopic groups in proteins

Michael P. Williamson[1] and Tetsuo Asakura[2]

[a]*Krebs Institute, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2UH, UK* and [b]*Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184, Japan*

Received 14 February 1992; revised version received 23 March 1992

We have calculated chemical shifts for a range of diastereotopic protons in proteins (i.e. methylene protons, and the methyl groups of valine and leucine residues), using a recently optimised method for chemical shift calculation. The calculations are based on crystal structure coordinates, and have been compared with experimental stereospecific assignments. The results indicate that chemical shifts can be used to suggest stereospecific assignments with about 80% probability of being correct, in cases where both the experimental and the calculated chemical shift differences between a pair of diastereotopic protons are greater than 0.3 ppm. Inaccurate calculations are shown to be caused in most cases by differences between crystal and solution structures. Furthermore, chemical shift calculations based on NMR structures are shown to be capable of acting as a further constraint on structure, by limiting the range of side-chain conformations adopted in structures calculated from NMR data.

NMR; Chemical shift; Stereospecific assignment; Protein structure

1. INTRODUCTION

As methods for the calculation of protein structures from NMR data have developed, it has become clear that the inclusion of constraints from stereospecifically assigned groups (namely, methylene protons and the methyl groups of valine and leucine) is crucial for the improvement of the precision of structures. Although some stereospecific assignments have been made using isotopic labelling [1,2], this is not a generally applicable method, and most such assignments are made by consideration of coupling constants and nuclear Overhauser effects (NOEs) [3]. These methods are usually secure and unambiguous, but it is desirable to have an independent way of checking the stereospecific assignments, particularly since incorrect stereospecific assignments can be hard to spot, and can have the effect of improving the apparent precision of structures, while lowering the accuracy [4]. The aim of this Letter is to study whether chemical shifts can be used for this purpose, and further to show how chemical shifts of diastereotopic groups can be used in the refinement of structures generated from NMR data.

2. MATERIALS AND METHODS

Chemical shifts were calculated from structures in the Brookhaven data base, namely 1R69, 1CSE, 1LZT, 1HOE, 2TRX, and 2A1T, for the structures of 434 repressor, eglin c (in a complex with subtilisin Carlsberg), hen egg-white lysozyme, tendamistat (α-amylase inhibitor HOE-467A), *E. coli* thioredoxin, and tendamistat (9 NMR structures). The structure of the squash seed trypsin inhibitor CMTI-I was kindly provided by Dr. W. Bode (Max-Planck-Institut, Martinsried), and the structure of barnase by Prof. G.G. Dodson (York). Protons were added to the structures using standard geometries. For methyl groups, protons were added in staggered geometry, and chemical shifts were calculated by averaging over the three positions. Chemical shifts were calculated using the program SHIFTCALC, described more fully elsewhere [5]. Chemical shifts are calculated as

$$\sigma = \sigma^{ring} + \sigma^{le} + \sigma^{ani}$$

where σ^{ring} is the ring-current chemical shift, calculated using the Johnson–Bovey algorithm, σ^{le} is the electric field shift, caused by polarisation of the C–H bond by point atomic charges on atoms C′, O, N and HN, and σ^{ani} is the magnetic anisotropy shift, caused by the anisotropy of C′=O and C′–N bonds in the peptide groups in backbone and side chain. When calculating shifts of the C$^\beta$H protons of aromatic residues, the σ^{ring} calculation included the shifts caused by intraresidue ring-current shifts.

Torsion angles were measured using a routine of the program FRODO. Standard deviations were calculated as $[\Sigma(x_i - \langle x_i \rangle)^2/(n-1)]^{1/2}$. Stereospecific assignments were taken from the literature [1,6–9], except for those of thioredoxin and barnase, which were kindly provided by David LeMaster and Mark Byeroft, respectively. Stereospecific assignments for methylene protons in eglin c were modified from those published [6] in the light of recent results (M.S. Goldberg, S.G. Hyberts and G. Wagner, to be published).

3. RESULTS

Chemical shifts were calculated for all the diastereotopic groups in proteins where stereospecific assignments have been published, and where the crystal structure is available. From these results we calculated the difference between the chemical shifts for the two dia-

Correspondence address: M.P. Williamson, Krebs Institute, Dept. of Molecular Biology and Biotechnology, P.O. Box 594, University of Sheffield, Western Bank, Sheffield S10 2UH, UK. Fax: (44) (742) 728697.

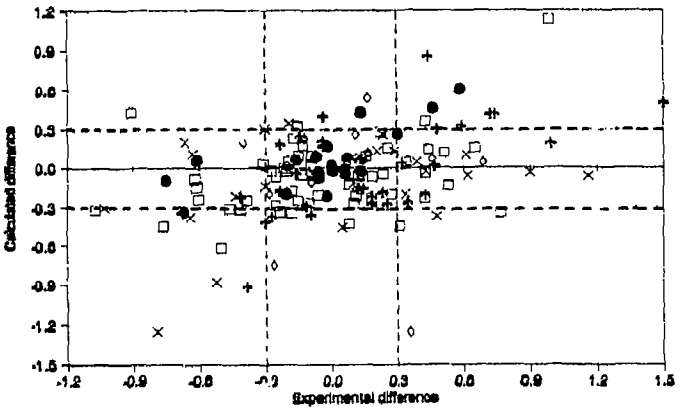

Fig. 1. Difference in chemical shift between pairs of diastereotopic protons, expressed as (HA2–HA1), (HB3–HA2), etc. The plot compares calculated and experimental values for 434 repressor and thioredoxin (●, these assignments were made using isotopic labels), hen lysozyme (×), squash seed trypsin inhibitor and barnase (□), eglin c (◇), and tendamistat (+). The dashed lines mark the 0.3 ppm cutoffs used in this work. Only protons with both experimental and calculated shifts greater than 0.3 ppm are considered here.

stereotopic protons or groups, expressed as the difference HA2–HA1 for glycine, HB3–HB2 for $C^{\beta}H$, etc. This provides a single signed measure of the difference in chemical shift between the two groups. The difference was then compared to the experimental value. The results are shown in Fig. 1. Perfect calculations would show a perfect correlation running from bottom left to top right. The correlation is indeed present, but is weak.

In our previous calculations of chemical shifts of $C^{\alpha}H$ protons [5,10,11], we found that chemical shifts can be calculated with a standard deviation of around 0.3 ppm or better. As part of a preliminary study of diastereotopic protons [10], we suggested that useful consideration of chemical shifts of the diastereotopic glycyl $C^{\alpha}H$ protons should be limited to pairs of protons where the difference in chemical shift is greater than 0.3 ppm in both the calculated and the experimental data. For the sign of the calculated chemical shift to differ from that of the experimental shift in such cases, the calculated shift would have to be in error by at least 0.6 ppm or two standard deviations. In a normal distribution, the chance of this happening is less than 2.5%. In Table I, we have listed all those pairs of protons with calculated and experimental shift differences greater than 0.3 ppm, and we have indicated whether the sign of the shift is the same or different. The relevant protons are to be found in the outer quadrants of Fig. 1. There are 28 proton pairs with both experimental and calculated shift differences greater than 0.3 ppm. Of these, 23 have the same sign for both calculated and experimental shift, and may therefore be considered 'correct', while 5 have the opposite sign, and are designated 'incorrect'. The most outstandingly 'incorrect' result is that for Val^{34} of eglin c (experimental +0.36 ppm, calculated −1.24 ppm). This result should be treated as doubtful, since in the crystal structure the valine methyls are very close to an aromatic ring, and the covalent geometry of Val^{34} is distorted, so that the CG methyl carbons are almost coplanar with the CA–CB bond.

The second objective of this work was to investigate the application of chemical shift calculations to the refinement of protein structures generated from NMR data. To this end we calculated the chemical shifts for the diastereotopic protons of tendamistat listed in Table I, using the nine NMR structures for tendamistat [8] deposited in the Brookhaven data bank. These values are listed in Table II. We also tabulate the local conformations close to these protons for each of the structures, described by the first sidechain torsion angle χ_1.

4. DISCUSSION

The results listed in Table I and shown in Fig. 1 demonstrate that most stereospecific assignments falling within the criteria listed above would be made correctly using chemical shift calculations. Nevertheless, the results are less impressive than might be expected, based on results obtained for $C^{\alpha}H$ protons [5]. Only 82% of proton pairs with both experimental and calcu-

Table I

Diastereotopic protons with calculated and experimental shift differences greater than 0.3 ppm

Protein	Protons	Experimental	Calculated	Assignment[a]
434 repressor	Val^{54}HG	0.58	0.61	Y
CMTI-I[b]	Cys^{16}HB	0.31	−0.45	N
CMTI-I[b]	Cys^{22}HB	0.43	0.36	Y
Barnase	Asn^{23}HB	0.77	−0.34	N
Barnase	Asn^{41}HB	−1.07	−0.32	Y
Barnase	Lys^{49}HB	−0.46	−0.31	Y
Barnase	Leu^{63}HB	−0.50	−0.61	Y
Barnase	Trp^{71}HB	−0.41	−0.31	Y
Barnase	Arg^{72}HB	0.98	1.13	Y
Barnase	Asp^{75}HB	−0.76	−0.45	Y
Barnase	His^{102}HB	−0.91	0.43	N
Eglin c	Val^{34}HG	0.36	−1.25	N
Hen lysozyme	Phe^{3}HB	0.48	−0.37	N
Hen lysozyme	Tyr^{23}HB	−0.79	−1.25	Y
Hen lysozyme	Asn^{27}HB	−0.52	−0.88	Y
Hen lysozyme	Asp^{52}HB	−0.64	−0.38	Y
Hen lysozyme	Asp^{66}HB	−1.03	−0.30	Y
Tendamistat	Leu^{14}HB	0.59	0.32	Y
Tendamistat	Glu^{16}HB	0.74	0.41	Y
Tendamistat	Asp^{24}HB	0.72	0.41	Y
Tendamistat	Asn^{25}HB	1.50	0.50	Y
Tendamistat	Val^{35}HG	0.44	0.85	Y
Tendamistat	Val^{36}HG	−0.38	−0.92	Y
Tendamistat	Leu^{44}HB	−0.30	−0.42	Y
Tendamistat	Pro^{50}HB	0.48	0.30	Y
Tendamistat	Leu^{70}HB	−0.68	−0.34	Y
Thioredoxin	Gly^{51}HA	−0.67	−0.33	Y
Thioredoxin	Gly^{74}HA	0.46	0.46	Y

[a]Indicates whether the stereospecific assignment would have been made correctly using chemical shifts: Y=yes, N=no. [b]Squash seed trypsin inhibitor.

Table II

χ_1 angles and calculated shift differences HB3–HB2 for the 9 tendamistat NMR structures

Res.		Structure										
		Xtal	1	2	3	4	5	6	7	8	9	Mean ± S.D.[a]
14	χ_1	164	163	160	162	154	169	168	161	171	174	165 ± 6
	$\Delta\delta$	0.34	0.15	0.20	0.22	0.19	0.32	0.31	0.26	0.31	0.16	0.24 ± 0.07
16	χ_1	−162	172	160	164	163	161	170	161	176	174	167 ± 6
	$\Delta\delta$	0.41	0.49	0.33	0.52	0.42	0.52	0.42	0.31	0.39	0.36	0.42 ± 0.08
24	χ_1	−172	173	156	−37	156	157	154	173	150	153	137 ± 66
	$\Delta\delta$	0.40	0.32	0.35	−0.16	0.40	0.22	0.40	0.29	0.19	0.19	0.25 ± 0.17
25	χ_1	176	179	173	−176	−175	−170	174	−159	−176	−157	−176 ± 11
	$\Delta\delta$	0.47	1.02	1.02	1.00	1.03	0.88	0.99	0.67	0.89	0.68	0.91 ± 0.14
35	χ_1	178	106	117	107	108	122	106	106	107	107	110 ± 6
	$\Delta\delta$	0.84	−0.05	0.47	0.18	0.07	0.75	−0.14	0.34	−0.08	0.43	0.22 ± 0.30
36	χ_1	−176	−171	172	−173	178	173	−169	177	175	172	179 ± 8
	$\Delta\delta$	−0.93	0.24	0.61	0.38	0.34	0.35	−0.34	0.24	0.14	0.27	0.25 ± 0.26
44	χ_1	−66	−102	−55	−55	−55	−55	−56	−55	−99	−55	−65 ± 20
	$\Delta\delta$	−0.41	−0.38	0.88	0.25	0.62	0.14	0.14	0.37	−0.37	−0.49	0.11 ± 0.50
70	χ_1	−67	−54	−57	−72	−60	−61	−60	−75	−85	−80	−67 ± 11
	$\Delta\delta$	−0.34	−0.28	−0.22	−0.20	−0.25	−0.22	−0.23	−0.22	−0.19	−0.20	−0.22 ± 0.03

[a]The mean and standard deviation refer to the 9 NMR structures only, and do not include the crystal structure.

lated shifts greater than 0.3 ppm are calculated correctly, compared with the 98% expected from a statistical analysis of $C^{\alpha}H$ shifts. There are several possible explanations for this larger than expected error. It could be a simple statistical fluctuation, it could be because some of the stereospecific assignments are incorrect, or it could be because the solution and crystal structures are different. While either of the first two is possible (and the second would merit further study), the third is a much more likely explanation. This can be illustrated by consideration of the crystal structure of barnase.

There are three molecules in the asymmetric unit in barnase, which have been refined independently. Comparison of local regions of the structure in the three molecules therefore indicates the extent of structural variability in the crystal, and provides an indication of places where calculations based on the crystal structure may be in error. Since our calculations are of $C^{\beta}H$ protons, we consider variations in the χ_1 angle in the three structures. These are listed in Table III. All of the χ_1 angles are well maintained in the three molecules, except for the χ_1 angle of His^{102}. The χ_1 angles for which the side chain adopts the low energy staggered rotamers are −60, +60 and 180°, and the angles in the crystal structure suggest that at least two of these minima are likely to be populated in solution. This may account for the fact that the HB shift difference for His^{102} is incorrectly calculated.

The other diastereotopic pair incorrectly calculated in barnase is that of Asn^{23} HB. The orientation of the side chain of Asn^{23} is well maintained in the crystal, and variability of the side-chain position is thus not a good

Table III

χ_1 angles in the barnase crystal structures.

Molecule	Residue							
	23	41	49	63	71	72	75	102
A	54	70	−48	−64	−70	−85	−81	−86
B	54	71	−59	−75	−66	−75	−73	−99
C	55	68	−56	−55	−73	−84	−78	−135
Mean	54	70	−54	−65	−70	−81	−77	−107
± S.D.	1	2	6	10	4	6	4	25

explanation of the poor result. However, a possible explanation lies in the observation that the side chain of Asn^{23} is adjacent to the N terminus of the protein, which is disordered in the crystal; we were therefore unable to calculate the chemical shift effect on Asn^{23} arising from the first two residues (three in structure C), since there are no coordinates for these residues. We conclude that chemical shift calculations of diastereotopic protons can only be used as a guide to stereospecific assignments when the solution and crystal structures are closely identical.

Turning now to the use of chemical shift calculations for the refinement of NMR structures, the position is clearer. As might be expected, there is a large degree of variation in the chemical shifts calculated for the 9 tendamistat NMR structures (Table II), although the variability in χ_1 angles is lower. In some cases, the χ_1 angles appear to follow a bimodal distribution. Thus, for Asp^{24}, χ_1 is around +160° for eight structures and is −37° for structure 3. Either range is compatible with the input constraints, which included a restriction on χ_1 to the ranges 120 to 180° or −180 to 0° [8]. Similarly, for Leu^{44}, χ_1 is either around −55° or around −100°; both are within the χ_1 constraint of −120 to 0°. On inspection of the chemical shifts calculated for the $C^{\beta}H$ of Asp^{24} in the NMR structures, it is seen that most structures give calculated differences HB3−HB2 of +0.30 ± 0.09 ppm, compared to a difference calculated from the crystal structure of +0.40 ppm, and an experimental difference of 0.72 ppm. The exception is structure 3, with a calculated shift difference of −0.16 ppm. This figure is sufficiently different from the experimental result to imply that the χ_1 angle of −37° found for this structure is not compatible with the observed chemical shifts, although it presumably is compatible with NOE and coupling constant constraints. Residue 44 presents a more complicated case, as the chemical shift difference is strongly influenced by other factors than just χ_1. Nevertheless, the structures that best fit the observed difference (−0.30 ppm) tend to be those with a more negative χ_1, suggesting that the dominant χ_1 found in most NMR structures, of about −55°, is less negative than the true average value in solution. In agreement with this, the angle found in the crystal is more negative than −55°, at −66°.

We therefore conclude that chemical shift differences between diastereotopic proton pairs can be used as a further parameter to limit the range of conformations calculated using NMR data. Chemical shift calculations are strongly dependent on the exact geometry of the surrounding groups [10], and therefore are only applicable as a structure refinement tool at a late stage in the structure calculation, after NOE and coupling constant restraints have been applied. More general applications of chemical shift-based structure refinement are under investigation.

Acknowledgements: We thank Prof. Wagner, Dr. Bycroft and Dr. LeMaster for providing unpublished stereospecific assignments, and Dr. Bode for the crystal coordinates of CMTI-I.

REFERENCES

[1] Neri, D., Otting, G. and Wüthrich, K. (1990) Tetrahedron 46, 3287–3296.
[2] LeMaster, D.M. and Richards, F.M. (1988) Biochemistry 27, 142–150.
[3] Wagner, G. (1990) Progr. Nucl. Magn. Reson. Spectrosc. 22, 101–139.
[4] Havel, T.F. (1991) Progr. Biophys. Molec. Biol. 56, 43–78.
[5] Williamson, M.P. and Asakura, T., submitted for publication.
[6] Hyberts, S.G., Märki, W. and Wagner, G. (1987) Eur. J. Biochem. 164, 625–635.
[7] Smith, L.J., Sutcliffe, M.J., Redfield, C. and Dobson, C.M. (1991) Biochemistry 30, 986–996.
[8] Kline, A.D., Braun, W. and Wüthrich, K. (1988) J. Mol. Biol. 204, 675–724.
[9] Holak, T.A., Gondol, D., Otlewski, J. and Wilusz, T. (1989) J. Mol. Biol. 210, 635–648.
[10] Williamson, M.P., Asakura, T., Nakamura, E. and Demura, M. (1992) J. Biomol. NMR 2, 83–98.
[11] Williamson, M.P. and Asakura, T. (1991) J. Magn. Reson. 94, 557–562.