*Journal of Biomolecular NMR*, **13:** 199–211, 1999.
*KLUWER/ESCOM*



# $C^{\alpha}$ and $C^{\beta}$ carbon-13 chemical shifts in proteins from an empirical database

Mitsuo Iwadate[a], Tetsuo Asakura[a] & Michael P. Williamson[b,*]
[a]*Department of Biotechnology, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184, Japan;* [b]*Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, U.K.*

Received 14 September 1998; Accepted 5 November 1998

*Key words:* carbon-13 chemical shifts, hydrogen bonding, protein structure, structure refinement

**Abstract**

We have constructed an extensive database of $^{13}C$ $C^{\alpha}$ and $C^{\beta}$ chemical shifts in proteins of solution, for proteins of which a high-resolution crystal structure exists, and for which the crystal structure has been shown to be essentially identical to the solution structure. There is no systematic effect of temperature, reference compound, or pH on reported shifts, but there appear to be differences in reported shifts arising from referencing differences of up to 4.2 ppm. The major factor affecting chemical shifts is the backbone geometry, which causes differences of ca. 4 ppm between typical α-helix and β-sheet geometries for $C^{\alpha}$, and of ca. 2 ppm for $C^{\beta}$. The side-chain dihedral angle $\chi^1$ has an effect of up to 0.5 ppm on the $C^{\alpha}$ shift, particularly for amino acids with branched side-chains at $C^{\beta}$. Hydrogen bonding to main-chain atoms has an effect of up to 0.9 ppm, which depends on the main-chain conformation. The sequence of the protein and ring-current shifts from aromatic rings have an insignificant effect (except for residues following proline). There are significant differences between different amino acid types in the backbone geometry dependence; the amino acids can be grouped together into five different groups with different $\phi,\psi$ shielding surfaces. The overall fit of individual residues to a single non-residue-specific surface, incorporating the effects of hydrogen bonding and $\chi^1$ angle, is 0.96 ppm for both $C^{\alpha}$ and $C^{\beta}$. The results from this study are broadly similar to those from ab initio studies, but there are some differences which could merit further attention.

## Introduction

In the last few years there has been a resurgence in interest in understanding chemical shifts in proteins (Szilágyi, 1995; Williamson and Asakura, 1997; Ando et al., 1998a). This has stemmed largely from the large amounts of data now available, which have allowed earlier predictions to be tested and refined. There have been empirical studies of $^{1}H$ (Ösapay and Case, 1991; Williamson and Asakura, 1993), $^{13}C$ (Spera and Bax, 1991; Wishart and Sykes, 1994b) and $^{15}N$ (Le and Oldfield, 1994) shifts as well as a variety of ab initio approaches (de Dios, 1996; Sitkoff and Case, 1997; Ando et al., 1998b). These studies have suggested that there is a highly significant difference between the isotropic chemical shifts in α-helix and β-sheet of both $C^{\alpha}$ and $C^{\beta}$, which has stimulated the application of $^{13}C$ shifts to the determination of secondary structure in proteins through the Chemical Shift Index (Wishart and Sykes, 1994a), and to the direct refinement of protein structures against $^{13}C$ shifts (Laws et al., 1993; Le et al., 1995; Pearson et al., 1995; Oldfield, 1995; Kuszewski et al., 1995; Luginbühl et al., 1995; Beger and Bolton, 1997; Clore and Gronenborn, 1998).

However, despite these developments, there remain many uncertainties in the relationship between protein structure and $^{13}C$ shift. The empirical stud-

*To whom correspondence should be addressed. E-mail: M.Williamson@Sheffield.ac.uk.

*Supporting information available*: The first page of the database (1 page: the entire database is available electronically from the Websites quoted in this paper); tables of the effect of backbone hydrogen bonding on $C^{\alpha}$ and $C^{\beta}$ shifts, in 20° × 20° regions (4 pages); a summary of the effect of following and previous residue on $C^{\alpha}$ and $C^{\beta}$ shifts (1 page).

ies have been based on a fairly small database, and the ab initio studies are of necessity based on small peptide fragments, which begs the question of how reliable the calculations are when applied to proteins. Therefore we felt it to be important to assemble a larger database of $^{13}C$ shifts in proteins of known structure, to enable us to study the effect of protein conformation and sequence on $C^{\alpha}$ and $C^{\beta}$ shifts. These results in turn will help to guide efforts in protein structure refinement using $^{13}C$ shifts.

## Methods

The protein coordinates were taken from crystal structures of proteins in the Brookhaven Protein Data Bank (Bernstein et al., 1977), obtained at high resolution, and for which NMR studies have indicated that the solution structure is essentially identical to the crystal structure. Secondary structure in the proteins was determined by the method of Kabsch and Sander as implemented in the program ksdssp (MidasPlus 2.0, Computer Graphics Laboratory, University of California, San Francisco, CA), and hydrogen bonding was assessed by a locally written algorithm based on the method of Kabsch and Sander (1983).

For each amino acid type, shift data were converted to secondary shifts by subtracting random coil shift values, using the values of Wishart et al. (1995a). No data were included for cysteine/cystine residues or for histidine residues, because random coil values for these residues are uncertain and may depend on protonation state. To construct shielding maps, showing how the shift varies with the backbone dihedral angles $\phi$ and $\psi$, the method of Spera and Bax (1991) was used: the individual secondary shift values $\delta(\phi_k, \psi_k)$ were convoluted with a Gaussian function prior to addition and normalisation to obtain smoothed surface grid points $\Delta(\phi,\psi)$:

$$\Delta(\phi, \psi) = \frac{\sum \delta(\phi_k, \psi_k) \exp(-((\phi - \phi_k)^2 + (\psi - \psi_k)^2)/450)}{\sum \exp(-((\phi - \phi_k)^2 + (\psi - \psi_k)^2)/450}$$

Shielding surface values were only considered significant in regions of the plot where the residue density function, $\Sigma_k \exp(-((\phi - \phi_k)^2 + ((\psi - \psi_k)^2)/450)$, is larger than 3.1, where the summation extends over all residues *k*.

Shielding surfaces were calculated for each protein separately. The typical standard deviation of observed chemical shifts $\delta(\phi_k, \psi_k)$ from the smoothed surface was just greater than 1 ppm. Amino acids whose secondary shift differed from the corresponding point on the surface by more than 2 ppm were excluded, on the grounds that it is likely that the residue in question either has a different conformation in solution from that in the crystal, or has an error in the published shift. The proportion of residues excluded on this criterion was 8.8% for $C^{\alpha}$ and 8.5% for $C^{\beta}$ (367 $C^{\alpha}$ and 259 $C^{\beta}$). Many of the residues excluded in this way were in loops or at chain termini, supporting this justification. In addition, two proteins (lambda repressor and c-H-ras) were omitted entirely from the list, because they had a large number of residues with shifts that were more than 2 ppm from the surface. One of these (c-H-ras) has an alternative assignment table which proved suitable for inclusion in the database (cf. Table 1).

After shielding surfaces had been calculated for each protein, and individual residues had been excluded as described above, the data from all proteins were combined and an overall shielding surface was calculated. Data from each protein were then corrected, to account for likely inconsistencies in the method of referencing of data from different proteins. For each protein, a correction was calculated that, when added to the $C^{\alpha}$ and $C^{\beta}$ shifts, gave the smallest root-mean-square difference between the data for the protein and the overall shielding surface. The correction was applied to each protein, after which the overall shielding surface and any further corrections were re-calculated. This second iteration resulted in very small changes to both the surface and individual correction values.

To assess the effect of hydrogen bonding on chemical shifts (which might be expected to depend strongly on the backbone conformation), residues were grouped into $20^{\circ} \times 20^{\circ}$ regions, and within each region were placed into one of four bins: C+N+, C+N−, C−N+, or C−N−, where C+/− indicates that the carbonyl oxygen is/is not hydrogen bonded, and N+/− indicates that the backbone nitrogen is/is not hydrogen bonded. A Student's *t*-test was used to test whether there is any significant change in chemical shift (at the 5% level) on hydrogen bonding to carbon (i.e. comparing C+N+ and C+N− to C−N+ and C−N−) and nitrogen (C+N+ and C−N+ versus C+N− and C−N−). Tests for $C^{\alpha}$ and $C^{\beta}$ were carried out separately.

In the cross-validation approach to check the validity of hydrogen-bond corrections, the database was divided randomly into 10 equally sized groups. Ten sets of calculations were then conducted, in each of which one group of 10% was omitted. In each set of calculations, the corrections required for the three $\phi,\psi$ regions (as defined in Table 4) were calculated as described in the preceding paragraph. $C^{\alpha}$ and $C^{\beta}$ shielding surfaces were calculated for the remaining 90% of data, both without and with hydrogen bonding corrections applied.

The effect of sequence was analysed by considering pairs of residues X-Y, in which both X and Y were in a helix, or both X and Y were in a β-sheet, or there was no restriction on their conformation. For the effect of the following residue, X was kept fixed and the chemical shift effect of the 18 (i.e. all residues except His and Cys) different Y was analysed. This was repeated for all possible X, and the values were averaged. A similar analysis with fixed Y was used for the effect of the preceding residue.

Ring-current shifts were calculated using a standard Haigh–Mallion calculation (Haigh and Mallion, 1980) with ring intensity factors derived from empirical fitting against $^{1}H$ data (Williamson and Asakura, 1993). The program can be found at http://www.sheffield.ac.uk/∼mbb/nmr/resources.html.

## Results

### *Protein $^{13}C$ database*

The database contains 3796 $^{13}C^{\alpha}$ and 2794 $^{13}C^{\beta}$ shifts, from 40 different proteins, as described in Table 1. All the proteins have high-resolution crystal structures, and NMR studies have indicated that the solution conformation of the protein is essentially identical to that in the crystal. The dihedral angles $\phi$, $\psi$ and $\chi^1$ were calculated from the crystal structure, as were locations of regular secondary structure and the presence of any intramolecular hydrogen bonds to backbone atoms. The database is available at http://www.shef.ac.uk/∼mbb/nmr/resources.html or at http://www.tuat.ac.jp/∼asakura/research.

It has been previously noted that different $^{13}C$ assignments have used different methods for chemical shift referencing, and that the different standards, such as trimethyl silyl propionate (TSP), dioxane and tetramethyl silane (TMS) may give different shifts, which in addition will vary with temperature and pH (Wishart et al., 1995b). Furthermore, the reported reference standard may in many cases be a secondary standard, and thus may not in fact reflect the method used for referencing; it is now common to reference the $^{1}H$ spectrum using an easily observed standard, and reference the $^{13}C$ spectrum using the relative gyromagnetic ratios of $^{1}H$ and $^{13}C$ (Wishart et al., 1995b). Therefore, it was of interest to see whether there was any relationship between the reference compound, temperature and pH described for each protein, and the correction needed to bring the data for that protein into harmonisation with those for the others. The correction was calculated iteratively as described in the Methods section, and is listed in Table 1. It is significant that some proteins required corrections of more than 1 ppm, and that the total range of corrections was 4.2 ppm, with a root-mean-square correction of 0.77 ppm. There was no systematic effect of the reference compound used; as noted above, significant differences have been noted between the shifts obtained from different compounds (Wishart et al., 1995b), and this lack of effect may be due more to the fact that, although the majority of studies claimed that shifts were referenced to TSP, the way that this referencing was done may have differed significantly between studies. Adherence to the IUPAC recommendations for chemical shift referencing should help to reduce this problem in the future (Markley et al., 1998). There was also no significant effect of pH or temperature on the corrections needed (Figure 1).

### *Effect of backbone dihedral angle*

It has been well documented that the backbone dihedral angles $\phi$ and $\psi$ have a large effect on chemical shifts of $C^{\alpha}$ and $C^{\beta}$ atoms (Spera and Bax, 1991; Asakawa et al., 1994a; Wishart and Sykes, 1994b). At the outset of this study, the dihedral effect seemed likely to be the largest single factor controlling chemical shifts: fortunately, the results reported here confirm this. Thus it is appropriate to look first at the effect of backbone conformation, and treat other factors as minor perturbations to this major effect.

We show below that there are significant differences in the shielding maps of different amino acids. However, all amino acids have a basically similar behaviour, as shown in Figure 2. The results show that, as has previously been described, there is a marked difference between the shifts of $C^{\alpha}$ and $C^{\beta}$ in α-helical and β-sheet regions. However, the left-handed helix region is not greatly different from an area between helix and sheet, making it difficult to distinguish the two regions on the basis of chemical shift alone. The

*Table 1.* Proteins used to construct the database of $C^{\alpha}$ and $C^{\beta}$ shifts

| Protein | # Res | Assignment[a] | pH | T (°C) | Ref | PDB | Resolution (Å) | Correction[b] |
|---|---|---|---|---|---|---|---|---|
| Tendamistat | 74 | a | 3.2 | 47 | Dioxane | 1hoe | 2.00 | 0.71 |
| BPTI | 58 | b | 4.6 | 36 | TSP | 5pti | 1.70 | 1.30 |
| Interleukin 1β | 153 | c | 5.4 | 30 | TSP | 5i1b | 2.10 | 0.30 |
| Turkey ovomucoid 3 | 56 | d | 4.1 | 30 | Dioxane | 3sgb | 1.80 | −0.44 |
| Staphylococcal nuclease | 143 | e | 5.5 | 45 | TSP | 1snc | 1.65 | 1.44 |
| Ribonuclease H | 155 | f | 5.5 | 27 | TSP | 2rn2 | 1.48 | 1.70 |
| FK506 binding protein | 107 | g | 6.5 | 30 | TSP | 1fkf | 1.70 | −0.40 |
| Ferredoxin | 98 | h | 7.1 | 25 | Dioxane | 1fxa | 2.50 | −0.21 |
| Phosphocarrier III | 169 | i | 6.4 | 37 | TSP | 1f3g | 2.10 | −0.22 |
| Glucose permease IIA | 162 | j | 6.6 | 35 | TSP | 1gpr | 1.90 | 1.89 |
| Horse ferrocytochrome c | 104 | k | 5.5 | 27 | Dioxane | 2pcb | 2.80 | −0.34 |
| Flavodoxin | 169 | l | 7.5 | 26 | TMS | 1flv | 2.00 | 0.21 |
| Cyclophilin | 165 | m | 6.5 | 20 | TSP | 2rma | 2.10 | 0.34 |
| Thioredoxin | 108 | n | 5.7 | 35 | TMS | 2trx | 1.68 | −0.06 |
| Interleukin 4 | 133 | o | 5.7 | 36 | TSP | 1rcb | 2.25 | −0.21 |
| CheY | 128 | p | 6.5 | 30 | TSP | 1chn | 1.80 | 0.03 |
| Trp repressor | 108 | q | 6.0 | 45 | TSP | 2wrp | 1.65 | 0.18 |
| Pike parvalbumin | 109 | r | 5.7 | 32 | TSP | 1pva | 1.65 | −0.20 |
| Human profilin | 139 | s | 6.4 | 2 | TSP | 2btf | 2.55 | −0.37 |
| *Acanthamoeba* profilin | 125 | t | 6.5 | 30 | TSP | 1acf | 2.00 | −0.36 |
| Barstar | 89 | u | 6.7 | 37 | TSP | 1brs | 2.00 | 1.85 |
| hG-CSF | 172 | v | 3.5 | 27 | TMS | 1rhg | 2.20 | −0.31 |
| Heat shock TF | 88 | w | 3.4 | 25 | TSP | 2hts | 1.83 | −0.02 |
| HPr | 85 | x | 6.5 | 30 | TSP | 1poh | 2.00 | −0.05 |
| Cow tyrosine phosphatase | 157 | y | 5.0 | 30 | TSP | 1pnt | 2.20 | −0.15 |
| Chicken cystatin | 116 | z | 5.5 | 27 | TSP | 1cew | 2.00 | 0.43 |
| MCP-1 | 71 | aa | 5.4 | 35 | TSP | 1dok | 1.85 | 0.03 |
| Grb2 SH2 domain | 217 | bb | 6.2 | 25 | TSP | 1gri | 3.10 | 0.00 |
| Holo I-FABP | 131 | cc | 7.2 | 37 | TSP | 1icm | 1.50 | −0.78 |
| Apo I-FABP | 131 | cc | 7.2 | 33 | TSP | 1ifc | 1.19 | −0.63 |
| PGK | 394 | dd | 6.6 | 25 | TSP | 1php | 1.65 | −0.17 |
| Ubiquitin | 76 | ee | 5.7 | 30 | TMS | 1ubq | 1.80 | −2.29 |

overall standard deviation between the surface and the individual values used to construct the surface is 0.99 ppm for $C^{\alpha}$ and 0.97 ppm for $C^{\beta}$, using only regions where the density of data is sufficient to give a meaningful fit (using all the data gives values of 1.00 and 1.05 ppm, respectively). Thus, although there is statistically a big difference between α-helical and β-sheet regions, individual amino acids can have shifts that are so different from standard values that they can adopt values characteristic of the alternative secondary structure.

*Effect of side-chain geometry*

Ab initio studies from Oldfield's group have indicated that the chemical shifts of $C^{\alpha}$ and $C^{\beta}$ in valine residues can vary by almost 12 ppm, depending on the side-chain dihedral angle $\chi^1$; they have also shown that the $\chi^1$ dependence is somewhat different depending on the backbone geometry (Pearson et al., 1997; Havlin et al., 1997). The analysis is made simpler by the fact that most amino acids are strongly clustered into the three staggered rotamers *gauche*, *gauche*$^-$ and *trans* (Figure 3). If we consider only the ab initio shifts in the three staggered rotamers, we might expect $\chi^1$ differences of up to 2 ppm in $C^{\alpha}$ and 3 ppm in

*Table 1.* (continued)

| Protein | # Res | Assignment[a] | pH | T (°C) | Ref | PDB | Resolution (Å) | Correction[b] |
|---|---|---|---|---|---|---|---|---|
| Transforming growth factor | 112 | ff | 4.2 | 45 | TSP | 2tgi | 1.80 | −0.75 |
| Cytochrome c2 | 112 | gg | 8.0 | 35 | TSP | 3c2c | 1.68 | −0.22 |
| c-H-ras p21 | 169 | hh | 6.5 | 30 | TSP | 4q21 | 2.00 | −0.10 |
| Calmodulin | 148 | ii | 6.3 | 47 | TSP | 4cln | 2.20 | −0.38 |
| Recoverin | 198 | jj | 6.8 | 30 | TSP | 1rec | 1.90 | 0.06 |
| Ribonuclease T1 | 104 | kk | 5.5 | 40 | TSP | 9rnt | 1.50 | −0.72 |
| Phospholipase A2 | 123 | ll | 4.3 | 40 | TMS | 2bpp | 1.84 | −0.25 |
| Troponin C | 162 | mm | 7.0 | 40 | DSS | 1top | 1.78 | −0.33 |

[a]References: a, Kessler et al. (1990) *Biopolymers*, **30**, 465; b, Wagner et al. (1986) *Biochemistry*, **25**, 5839; c, Clore et al. (1990) *Biochemistry*, **29**, 8172; d, Robertson et al. (1990) *Biopolymers*, **29**, 461; e, Wang et al. (1990) *Biochemistry*, **29**, 102; f, Yamazaki et al. (1993) *Biochemistry*, **32**, 5656; g, Xu et al. (1993) *Biopolymers*, **33**, 535; h, Oh & Markley (1990) *Biochemistry*, **29**, 3993; i, Pelton et al. (1991) *Biochemistry*, **30**, 10043; j, Fairbrother et al. (1992) *Biochemistry*, **31**, 4413; k, Gao et al. (1990) *Eur. J. Biochem.*, **294**, 355; l, Stockman et al. (1990) *Biochemistry*, **29**, 9600; m, Neri et al. (1991) *FEBS Lett.*, **294**, 81; n, Chandrasekhar et al. (1994) *J. Biomol. NMR*, **4**, 411; o, Powers et al. (1992) *Biochemistry*, **31**, 4334; p, Moy et al. (1994) *Biochemistry*, **33**, 10731; q, Hong et al. (1994) *J. Mol. Biol.*, **238**, 592; r, Alattia et al. (1996) *Eur. J. Biochem.*, **237**, 561; s, Metzler et al. (1993) *Biochemistry*, **32**, 13818; t, Archer et al. (1993) *Biochemistry*, **32**, 6680; u, Lubienski et al. (1994) *Biochemistry*, **33**, 8866; v, Zink et al. (1992) *FEBS Lett.*, **314**, 435; w, Damberger et al. (1994) *Protein Sci.*, **3**, 1806; x, Nuland et al. (1992) *Eur. J. Biochem.*, **203**, 483; y, Logan et al. (1994) *Biochemistry*, **33**, 11087; z, Engh et al. (1993) *J. Mol. Biol.*, **234**, 1048; aa, Handel et al. (1996) *Biochemistry*, **35**, 6569; bb, Thornton et al. (1996) *Biochemistry*, **35**, 11852; cc, Hodsdon et al. (1997) *Biochemistry*, **36**, 1450; dd, Hosszu et al. (1997) *Biochemistry*, **36**, 333; ee, Wand et al. (1996) *Biochemistry*, **35**, 6116; ff, Hinck et al. (1996) *Biochemistry*, **35**, 8517; gg, Blanchard et al. (1997) *J. Biomol. NMR*, **9**, 389; hh, Campbell-Burk et al. (1992) *J. Biomol. NMR*, **2**, 639; ii, Ikura et al. (1990) *Biochemistry*, **29**, 4659; jj, Ikura et al. (1994) *Biochemistry*, **33**, 10743; kk, Pfeiffer et al. (1997) *J. Mol. Biol.*, **266**, 400; ll, Berg et al. (1995) *Nature Struct. Biol.*, **2**, 402; mm, Slupsky et al. (1995) *Protein Sci.*, **4**, 1279.
[b]Correction applied to published shifts to bring the data into closest agreement (see text).

$C^{\beta}$ (Havlin et al., 1997). The amount of data available for individual amino acids is very variable, as shown in Table 2. Thus, it is not possible to obtain enough data for some amino acids to clearly define chemical shift dependence on $\chi^1$, particularly if we wish to discriminate between amino acids in the α-helical and β-sheet regions. Therefore we clustered amino acids together into chemically similar groupings, having first checked from looking at the behaviour of individual amino acids that such a grouping was appropriate.

The result of this analysis is shown in Table 3. It can be seen from the table that there is no significantly useful difference between the results for residues in α-helical and β-sheet conformations, and hence we have only attempted to interpret the overall results for all backbone conformations. For long side-chains (i.e. lysine, arginine, glutamate, glutamine and methionine), there is no detectable effect on the chemical shift of $C^{\alpha}$ or $C^{\beta}$ arising from side-chain orientation. However, the other side-chains do show a small but significant effect, which is very similar for $C^{\alpha}$ and $C^{\beta}$, in that the *gauche* rotamer causes $C^{\alpha}$ and $C^{\beta}$ shifts to be negative (i.e. more shielded/upfield), by about 0.3 ppm compared to the other two rotamers. It is probably significant that the 'long side-chains' are the amino acid residues that are most likely to be disordered, and most likely to differ between solution and solid state. Therefore, it may be that the apparent lack of conformational effect on the 'long side-chain' group is due to imprecise data and conformational averaging: the real effects of side-chain conformation may be greater than suggested here.

*Effect of hydrogen bonding*

When considering the possible effect of hydrogen bonding on chemical shifts, it is important not to 'double count'. Thus, we have already noted a strong dependence of chemical shift on backbone conformation, and it is therefore important that any bias in sampling among different backbone conformations be properly taken into account so that it does not produce a misleading conclusion about hydrogen bonding. One way to do this is to restrict the analysis to small areas of ($\phi$ ,$\psi$) space, small enough that sampling bias is insignificant. The centre of the α-helical region is reasonably flat (Figure 2), and therefore one should

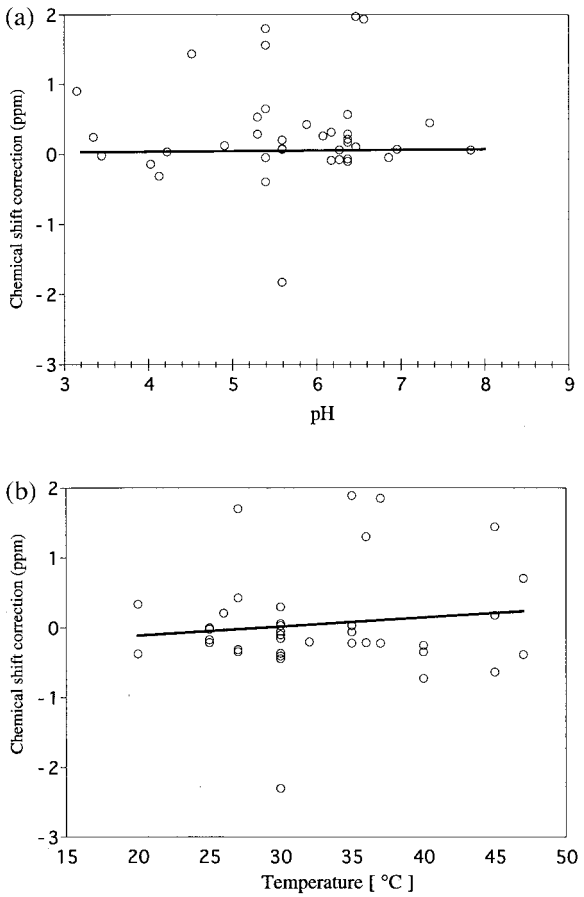


*Figure 1.* (a) Corrections to published shift needed to bring combined $C^{\alpha}$ and $C^{\beta}$ shifts into agreement, plotted against the pH at which the shifts are reported, for each of the 40 proteins listed in Table 1. The solid line is the line of best fit (correlation coefficient $R = 0.013$). (b) Shift corrections plotted against temperature. The solid line is the line of best fit ($R = 0.12$).

*Table 2.* Summary of data on individual amino acids

| Amino acid | # $C^{\alpha}$ shifts | # $C^{\beta}$ shifts |
|---|---|---|
| Ala | 309 | 246 |
| Arg | 165 | 117 |
| Asn | 179 | 149 |
| Asp | 245 | 195 |
| Gln | 143 | 119 |
| Glu | 321 | 246 |
| Gly | 336 | – |
| Ile | 206 | 161 |
| Leu | 344 | 239 |
| Lys | 318 | 263 |
| Met | 66 | 51 |
| Phe | 176 | 143 |
| Pro | 131 | 94 |
| Ser | 209 | 174 |
| Thr | 222 | 228 |
| Trp | 36 | 36 |
| Tyr | 118 | 96 |
| Val | 272 | 237 |
| Total | 3796 | 2794 |

be able to get a true idea of hydrogen bonding effects by restricting an analysis to this area. The result is shown in Figure 4: clearly, hydrogen bonding (as characterised by being in an α-helix or not) does lead to deshielding of $C^{\alpha}$. A further indication of the importance of hydrogen bonding comes from an analysis of how the $C^{\alpha}$ shift varies with position within an α-helix. Figure 5 shows that the $C^{\alpha}$ shifts of residues within α-helices change as the residue goes from being the C-terminal residue in a helix (position 0), and therefore with most likely no intramolecular hydrogen bond to the carbonyl, to a position inside the helix, and therefore with a carbonyl hydrogen bond. The mean secondary shift of residues at positions 0 and 1 is clearly more shielded than residues well within the helix, consistent with the conclusion above. It is however possible that some of this shift change may arise from partial unwinding of the helix, which would cause shifts to move towards their random coil values.

Strictly, from these results we can only draw such conclusions about residues in the α-helical conformation (as shown in Figure 4), because it could well be true that residues in other conformations have different behaviour on hydrogen bonding. Therefore, we carried out a more systematic survey of the effect of hydrogen bonding. As an example, the results for the effect of carbonyl hydrogen bonding on $C^{\alpha}$ shifts are shown in Figure 6. Similar results were obtained for hydrogen bonds to amide nitrogens, and for the effect on $C^{\beta}$ shifts; the results are available from our Web sites. If (as in Figure 6) amino acids are divided in 20° regions, there are not enough data to claim statistically meaningful effects for all but the most populated regions of $\phi,\psi$ space, as indicated by the figures in boldface in Figure 6. However, the data show a consistent pattern, in that they fall into three clearly defined regions, which can be identified roughly as α-helix, β-sheet, and a region in between. Within each region the same trends are apparent, as detailed in Table 4.

The validity of this approach was checked in two ways. In the first approach, hydrogen bonding corrections were calculated and tested using a jack-

*Table 3.* Dependence of chemical shift on side-chain rotamer (see Figure 3 for definition). Chemical shifts are given as difference from random coil ± standard deviation (number of amino acids used in the calculation). Only residues within 30° of the staggered rotamer minimum were considered in the analysis

| Amino acids | Rotamer | α-Region | β-Region | All residues |
|---|---|---|---|---|
| **(a) $C^{\alpha}$** | | | | |
| Val, Ile, Thr | $g^-$ | −0.46±0.77(25) | −0.25±0.84(56) | −0.30±0.87(178) |
| | *g* | −0.74±0.93(13) | 0.04±0.95(20) | −0.14±1.05(51) |
| | *t* | 0.23±0.84(100) | 0.07±0.87(189) | 0.15±0.92(377) |
| Phe, Tyr, Leu, Trp | $g^-$ | −0.18±0.99(99) | −0.02±0.82(99) | −0.15±0.94(316) |
| | *g* | −0.13±0.51(3) | −0.25±0.75(32) | −0.29±0.94(48) |
| | *t* | 0.20±0.77(109) | −0.02±0.80(55) | 0.18±0.89(201) |
| Lys, Arg, Glu, Gln, Met | $g^-$ | −0.08±0.86(200) | 0.13±0.77(68) | −0.10±0.87(412) |
| | *g* | 0.14±0.72(21) | −0.31±0.76(17) | −0.02±0.82(65) |
| | *t* | 0.21±0.80(125) | −0.23±0.65(70) | 0.04±0.84(252) |
| Asp, Asn | $g^-$ | 0.07±0.65(86) | 0.07±0.64(15) | 0.10±0.67(170) |
| | *g* | −0.25±0.83(6) | −0.25±0.46(6) | −0.42±0.66(81) |
| | *t* | 0.33±0.60(14) | −0.31±0.68(14) | −0.01±0.85(95) |
| **(b) $C^{\beta}$** | | | | |
| Val, Ile, Thr | $g^-$ | 0.20±0.82(22) | 0.26±0.81(52) | 0.09±1.00(162) |
| | *g* | 0.01±0.73(12) | −0.48±1.81(13) | −0.40±1.36(39) |
| | *t* | −0.07±0.61(113) | 0.06±1.01(161) | −0.01±0.88(341) |
| Phe, Tyr, Leu, Trp | $g^-$ | −0.06±0.99(74) | 0.30±0.94(77) | −0.03±1.02(243) |
| | *g* | −0.71(1) | −0.26±1.28(22) | −0.45±1.13(34) |
| | *t* | 0.09±0.82(82) | 0.23±1.10(38) | 0.10±0.91(148) |
| Lys, Arg, Glu, | $g^-$ | 0.03±0.70(150) | 0.17±1.11(57) | −0.02±0.91(323) |
| Gln, Met | *g* | 0.03±0.96(14) | 0.30±0.86(15) | −0.11±1.05(47) |
| | *t* | −0.03±0.66(95) | 0.14±0.97(47) | 0.01±0.83(186) |
| Asp, Asn | $g^-$ | 0.00±0.76(64) | 0.28±0.99(15) | 0.11±0.77(140) |
| | *g* | −0.94±0.81(7) | 0.12±0.91(4) | −0.30±0.78(68) |
| | *t* | 0.52±0.39(11) | 0.25±1.22(12) | 0.06±0.97(73) |

*Table 4.* Summary of hydrogen bonding effects on $C^{\alpha}$ and $C^{\beta}$, by region. The regions are defined in Figure 6

| Region | H-bond to NH | | H-bond to CO | |
|---|---|---|---|---|
| A | $C^{\alpha}$ +0.11 | $C^{\beta}$ −0.04 | $C^{\alpha}$ +0.44 | $C^{\beta}$ −0.14 |
| B | $C^{\alpha}$ −0.04 | $C^{\beta}$ +0.12 | $C^{\alpha}$ −0.20 | $C^{\beta}$ +0.19 |
| C | $C^{\alpha}$ −0.21 | $C^{\beta}$ +0.15 | $C^{\alpha}$ −0.13 | $C^{\beta}$ +0.19 |

knife/cross-validation method, in which 10% of the data are omitted and tested against the results from the optimised 90%. The results are listed in Table 4, and show that hydrogen bonding corrections range from −0.21 to +0.44 ppm. The standard error on these corrections, obtained from the jackknife procedure, is about 0.2 ppm. Simultaneous hydrogen bonds to both NH and CO give an increased effect in almost all cases, ranging from −0.34 to +0.55 ppm. When 90% of the data are used to construct a $\phi,\psi$ surface, the fit of the remaining 10% of data to the entire surface improves from 1.018 ± 0.040 ppm ($C^{\alpha}$) and 1.071 ± 0.039 ppm ($C^{\beta}$) to 1.001 ± 0.041 and 1.065 ± 0.034 ppm, respectively. In the second approach, all data in the database for amino acids that are hydrogen bonded to the backbone NH and/or CO were corrected using the values listed in Table 4 (thus, for example, residues in region C that are hydrogen bonded at NH only were corrected by adding −0.21 ppm to the $C^{\alpha}$ shift and +0.15 ppm to the $C^{\beta}$ shift). The $\phi,\psi$ surface was then recalculated using these results, to give a surface with a similar overall shape to the original surface but differing in detail. The fit of individual residues to the surface again improves by a small but significant amount: the rmsd for $C^{\alpha}$ decreases from 0.99 to 0.97 ppm, and for $C^{\beta}$ from 0.97 to 0.96 ppm.

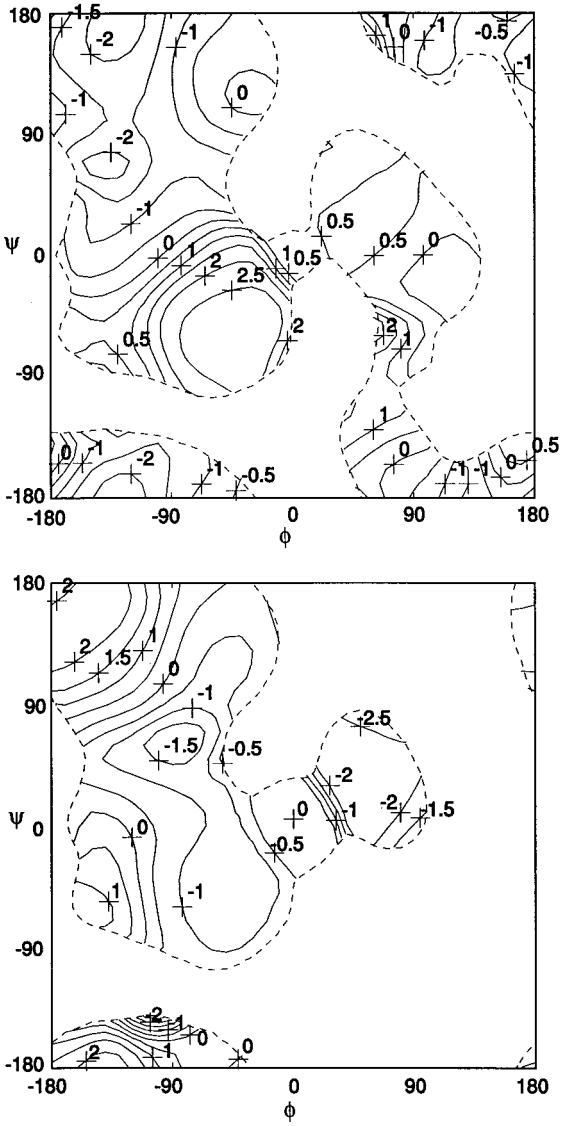


*Figure 2.* Smoothed shielding surface for all amino acids in the database. The contours show differences from the random coil shift, and the regions shown are those for which there are enough data to derive a reasonable surface. The two plots depict the shifts of (top) $C^{\alpha}$ and (bottom) $C^{\beta}$.

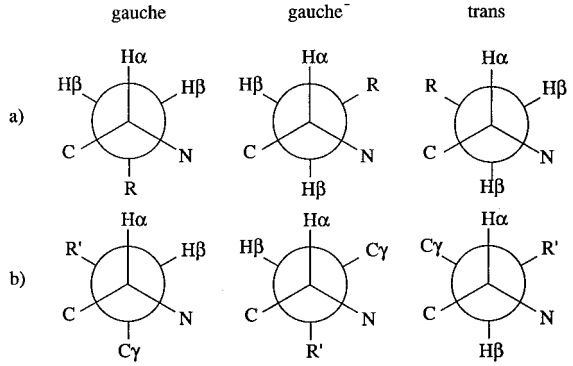


*Figure 3.* Definitions of the staggered rotamers of amino acids. Because of the way $\chi^1$ is defined, the same $\chi^1$ angle looks different for Val, Thr and Ile. (a) Amino acids with two $H^{\beta}$. The *gauche*, *gauche*$^-$ and *trans* rotamers correspond to $\chi^1$ angles of $+60°$, $-60°$ and $180°$, respectively. (b) Val, Thr and Ile, with only one $H^{\beta}$. For Val, the $\chi^1$ angles of the staggered rotamers are as above, but for Ile and Thr, the *gauche*, *gauche*$^-$ and *trans* rotamers correspond to $\chi^1$ angles of $180°$, $+60°$ and $-60°$, respectively.

### *Effect of sequence*

There is much evidence to suggest that the chemical shift of backbone $^{15}N$ atoms in proteins is dominated by sequence, i.e. it is strongly influenced by the nature of the preceding amino acid (Braun et al., 1994). It is probably also the case that carbonyl $^{13}C$ shifts are strongly influenced by the nature of the following residue, in that on denaturation much of the dispersion of $^{13}CO$ shifts remains (Yao et al., 1997). However, the dispersion of $C^{\alpha}$ and $C^{\beta}$ is largely lost on denaturation, suggesting that the chemical shifts of these nuclei are much less affected by sequence. It is possible that the effect of sequence depends on the local secondary structure. Therefore, sequence effects were analysed in a conformation-dependent manner. However, when sequence dependence in helical regions is compared to sequence dependence in extended regions, the results are almost indistinguishable, showing that there is in fact no obvious conformational dependence.

The results from our analysis (which can be found on the Web site) demonstrate that there is no significant effect of sequence on $C^{\alpha}$ and $C^{\beta}$ shifts: in almost all cases the neighbouring amino acid causes a change in chemical shift from random coil of less than 0.1 ppm, with a standard deviation of about 0.85 ppm. The only residue that had any possible effect is proline; when an amino acid is followed by proline, its $C^{\alpha}$ shift is altered by $-0.9 \pm 1.1$ ppm, and its $C^{\beta}$ shift is altered by $-0.5 \pm 1.2$ ppm. However, proline has a well-known propensity to change the conformation of the preceding amino acid, strongly encouraging a β-sheet-like or poly-proline II conformation (Williamson, 1994), and thus much of the chemical shift change could be accounted for by a conformational bias in the sampling of residues preceding proline compared to other amino acids. Therefore we can conclude that there is no evidence for any effect of amino acid sequence on shift.

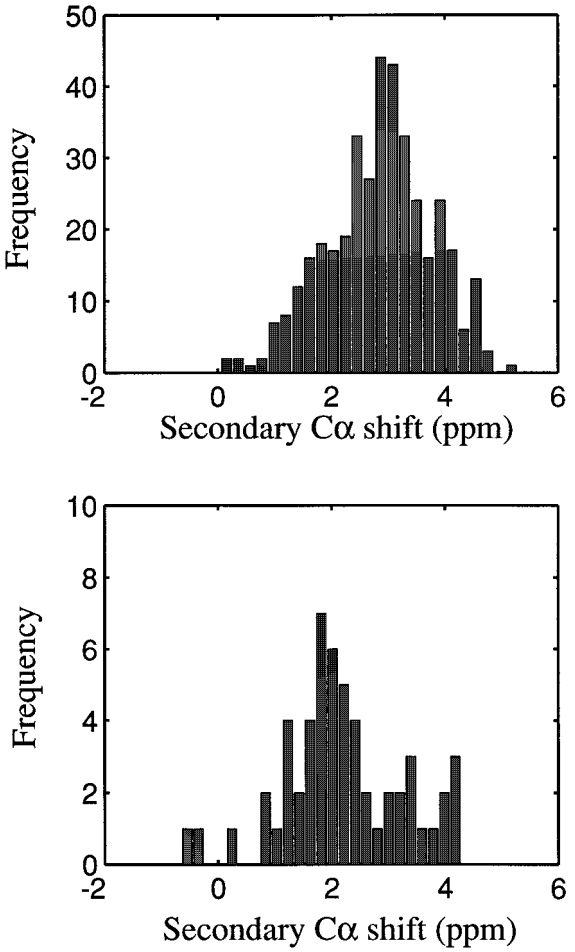


*Figure 4.* The effect of hydrogen bonding on $C^{\alpha}$ chemical shifts, for residues in helical conformation. Residues in the range $-60° < \phi < -20°$, $-80° < \psi < -10°$ were selected. (a) Residues in a regular α-helix, as identified by the Kabsch and Sander (1983) method. (b) Residues not in a regular α-helix. The chemical shifts are (a) $2.84 \pm 0.92$ ppm ($n = 372$) and (b) $2.19 \pm 1.07$ ppm ($n = 51$), which are very significantly different values as assessed by a Student's *t*-test ($p < 0.001$).

### *Effect of ring current*

A previous study has suggested that aromatic ring-current effects are generally small for $^{13}C$, except in a few individual cases, particularly close to large ring currents such as those from heme rings (Blanchard et al., 1997). This is because ring-current effects for $^{13}C$ and $^{1}H$ nuclei in the same geometry are equal in ppm, and are therefore much smaller for $^{13}C$ in comparison to its total chemical shift dispersion. This conclusion was borne out by the present study. There was no correlation between large deviations in data from the $\phi,\psi$ surface and calculated ring-current shifts. As a further test, all $^{13}C$ shifts were corrected by subtracting the

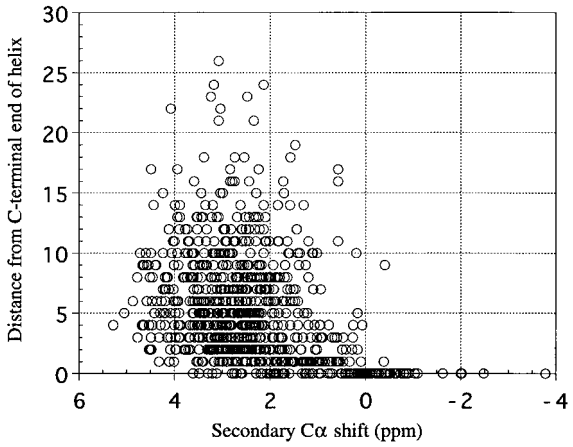


*Figure 5.* The influence of position within an α-helix on secondary shift. The position within a helix is shown as the distance in residues from the C-terminus of a helix, with position 0 being the C-terminal residue of the helix.

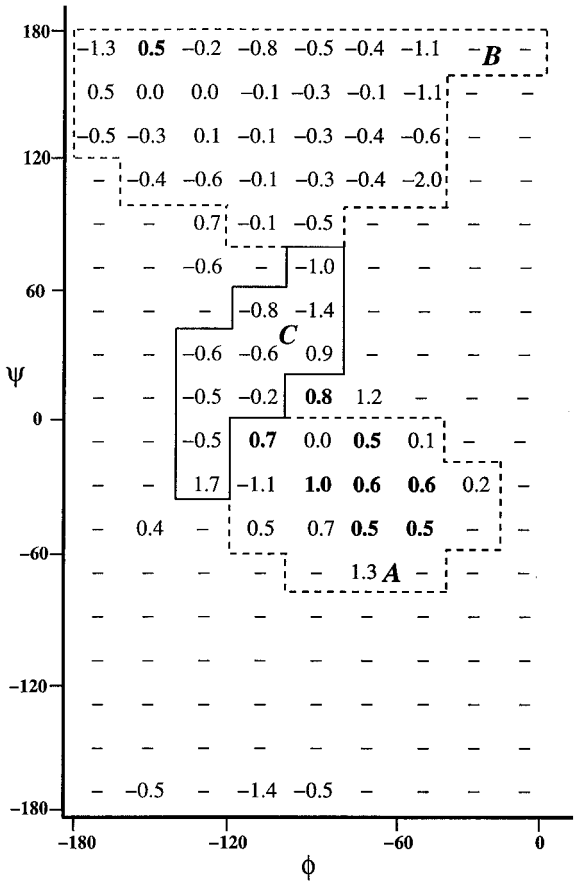


*Figure 6.* Effect on $C^{\alpha}$ shift of hydrogen bonding to the carbonyl group, as a function of backbone geometry. Results are grouped in 20° regions. A dash indicates that there are no data, and results in boldface are statistically significant at the 5% level. The dashed and solid lines indicate the three regions (A, B, C) identified in the text.

expected ring-current shift, and the surface was recalculated. There was almost no change in the overall rmsd to the surface.

*$\phi,\psi$ dependence of individual amino acids*

Theoretical calculations on alanine, glycine and valine have concluded that these different amino acids have different $\phi,\psi$ shift surfaces (Laws et al., 1993; Le et al., 1995; Havlin et al., 1997; Pearson et al., 1997). Empirical analysis from our database is difficult, because, although we have a large amount of data, the amount of data on each individual amino acid is limited. This means that the likely difference in shielding surface between different amino acids is probably of the same order as the random error in the surface, and therefore we cannot be certain that apparent differences between shielding surfaces for different amino acids are not due to random variability in the data. This conclusion is based on the observation that the similarity in shielding surface between different amino acids increases as we narrow the areas of comparison to areas with increasingly large densities of experimental data, up to a limiting residue density of about 4.0, after which further narrowing of the area of comparison gives no further increase in similarity.

We therefore concluded that meaningful comparisons of different amino acids could only be obtained by grouping together residues that have similar shielding surfaces. Shielding surfaces for the 18 amino acids (i.e. all the amino acids studied, omitting His and Cys) were clustered together, by picking clusters of amino acids with maximum similarity between shielding surfaces. The five clusters (in decreasing order of similarity) were: [Trp, Leu, Met, Phe, Tyr], [Glu, Gln, Lys, Arg, Asp, Asn], [Val, Ile, Thr, Pro], [Ala, Ser] and [Gly]. The results for each of these clusters are shown in Figures 7 and 8 for $C^\alpha$ and $C^\beta$ shifts, respectively.

## Discussion

Our initial assumption at the start of this study was that the largest variability in $C^\alpha$ and $C^\beta$ shifts in proteins comes from the backbone geometry, with smaller effects from other contributions such as side-chain orientation, hydrogen bonding etc. The results presented here have confirmed this: the backbone geometry can typically affect the $C^\alpha$ shift over a range of 5 ppm, while hydrogen bonding has a chemical shift range of 0.9 ppm, side-chain geometry has a range of almost 0.5 ppm, and sequence and ring-current shifts have negligible effects. $C^\beta$ shifts show a similar pattern. A previous shielding surface, constructed using four proteins, had an rms deviation between experimental and fitted data of 1.1 ppm (Spera and Bax, 1991). Our shielding surface, constructed using 40 proteins, has a similar appearance and has an rms deviation of 0.99 ppm for $C^\alpha$ and 0.97 ppm for $C^\beta$ (using data for the well-populated regions of $\phi,\psi$ space). When the effects of hydrogen bonding are included, by recalculating a modified surface for 'non-hydrogen bonded' residues and adding appropriate corrections for hydrogen bonding, the fit improves slightly to 0.97 and 0.96 ppm respectively. When hydrogen bonding and side-chain orientation effects are included simultaneously, the fit is again slightly improved, to 0.96 ppm for both $C^\alpha$ and $C^\beta$. A program for predicting $C^\alpha$ and $C^\beta$ shifts from the three-dimensional structure can be found on our Web site. Thus, the results presented here confirm and extend the previous study.

It is worth noting that the results presented here (particularly those on the effect of $\phi$ and $\psi$) are from a smoothed set of values, and also that the observed NMR parameters are necessarily obtained from proteins in solution, which are mobile. It may therefore be the case that the 'real' effects of structure on chemical shift are larger than those presented here. However, a comparison with ab initio results (see below) suggests that motional effects are not predominant.

We have compared our results to those obtained by ab initio calculations (Laws et al., 1993; Le et al., 1995; Havlin et al., 1997; Pearson et al., 1997). The general trends of the data are very similar, and improve in regions where the density of experimental data is greatest, suggesting that much of the apparent difference may be due to noise in the experimental results. For example, the ab initio calculations show that for $\phi = -90°$, the dependence of $C^\alpha$ shift on $\psi$ is much stronger for alanine than it is for glycine, in agreement with our experimental results. Our results suggest that there are real and significant differences between different amino acids, which will require considerably more experimental data to define satisfactorily. In view of the good agreement between experimental and calculated results, this implies that it would be useful to calculate shielding surfaces for other amino acid types. Both our results and the ab initio calculations (Havlin et al., 1997; Pearson et al., 1997) suggest that for valine, the *trans* side-chain rotamer produces deshielding of $C^\alpha$ shifts while the *gauche* and *gauche*$^-$ rotamers are roughly equally shielded. Our results sug-

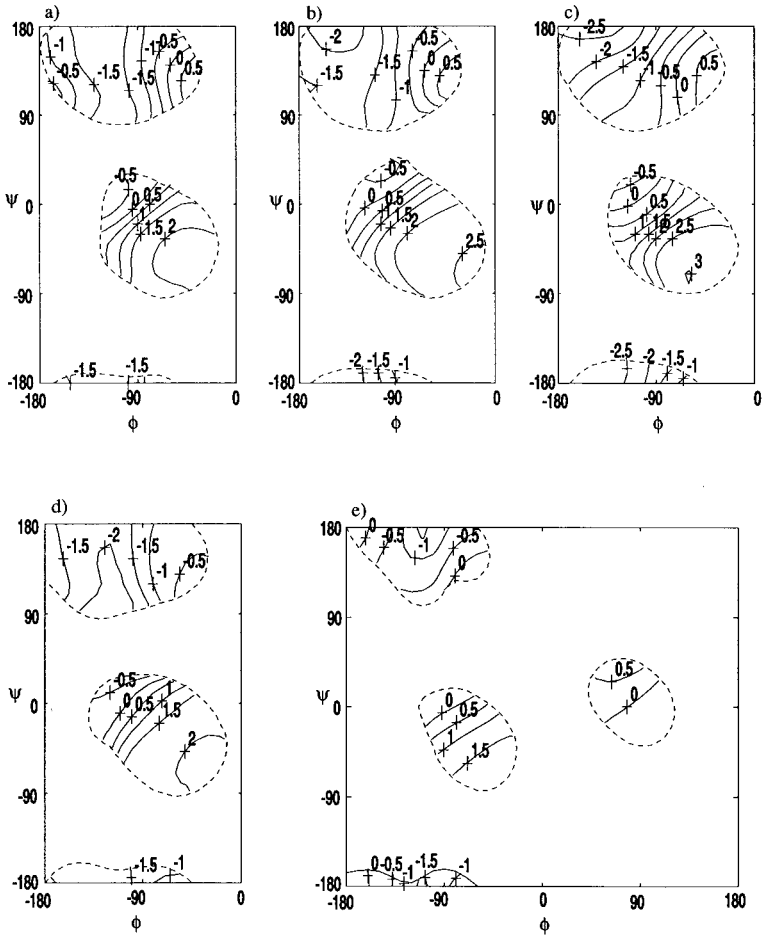


*Figure 7.* Shielding surface (smoothed chemical shift value as a function of $\phi$ and $\psi$) for $C^{\alpha}$, clustered into clusters of amino acid types with similar surfaces. (a) Trp, Leu, Met, Phe, Tyr; (b) Glu, Gln, Lys, Arg, Asp, Asn; (c) Val, Ile, Thr, Pro; (d) Ala, Ser; and (e) Gly. The surface is only shown for regions of the plot in which the residue surface density (i.e. the number of data points per degree$^2$) is greater than 3.1.

gest that this may be a more general phenomenon; however, it is difficult to determine how general the effect may be using experimental data, because they have many uncertainties, and therefore a more extended ab initio study is warranted. Theoretical studies on the effect of hydrogen bonding on $C^{\alpha}$ and $C^{\beta}$ shifts have been more equivocal, with one study (Asakawa et al., 1994b) concluding that hydrogen bonding may have an effect of similar magnitude to main-chain geometry on $C^{\alpha}$ shifts, but another (de Dios et al., 1994) concluding that the effect is only small. Our results show that hydrogen bonding has a small but non-negligible effect of up to 0.9 ppm, which is about five times smaller than the effect of main-chain geometry. Clearly, it would be useful for future calculations to address this question again.

This study provides a greatly improved database for the application of $C^{\alpha}$ and $C^{\beta}$ shifts to the refinement of protein structure, but it also demonstrates that we are currently unable to account for all the factors that govern individual $^{13}C$ shifts in proteins: the rms difference between predicted and experimental shifts is still uncomfortably high at around 0.96 ppm, with some individual values having much larger differences. Although some of these differences could be due to experimental error or to real differences between solution and crystal structures, it remains likely that we are currently unable to predict $^{13}C$ shifts in proteins to a better accuracy than 0.5 ppm. It therefore raises the question of what form $^{13}C$ chemical shift restraints should take. The success of the Chemical Shift Index (CSI) approach (Wishart and Sykes, 1994a) demonstrates that restraints derived from $^{13}C$

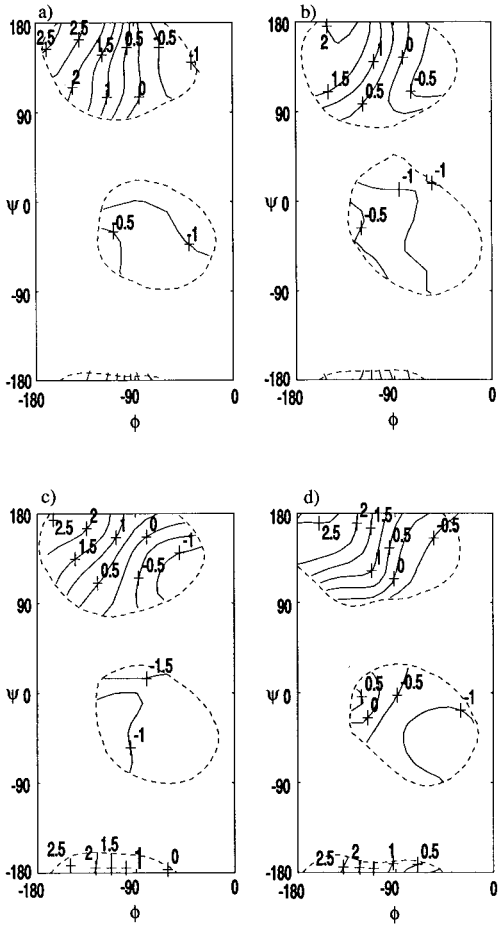


*Figure 8.* Shielding surface for $C^{\beta}$, clustered into clusters of amino acid types with similar surfaces. (a) Trp, Leu, Met, Phe, Tyr; (b) Glu, Gln, Lys, Arg, Asp, Asn; (c) Val, Ile, Thr, Pro; and (d) Ala, Ser. The surface is only shown for regions of the plot in which the residue surface density (i.e. the number of data points per degree$^2$) is greater than 3.1.

shifts are extremely useful. The power of the CSI is that $^{13}C$ shifts are used collectively to define regions of regular secondary structure, and therefore that individual anomalous values do not invalidate the method. The difficulty with direct refinement against $C^{\alpha}$ and $C^{\beta}$ shifts is that this will undoubtedly introduce a large number of incorrect, although weak, restraints (Clore and Gronenborn, 1998): even using a 1 ppm margin of error, with a standard deviation of 0.96 ppm and a normal distribution of errors, this means that statistically 30% of restraints may be incorrect. Direct chemical shift-based refinement has shown most promise in regions where other NMR-derived restraints are sparse, in particular in regions where there is no regular secondary structure, and in partially unfolded regions. It may therefore be that chemical shift refinement could be used in conjunction with conformational database potentials (Kuszewski et al., 1996) to generate conformational energy wells, which could direct peptide chains into the appropriate geometries. More work is required in this area.

## Acknowledgements

We thank Prof. Eric Oldfield for helpful comments. This work was supported by the International Human Frontier Science Program and the Program for Promotion of Basic Research Activities for Innovative Biosciences, Japan. M.P.W. is an Invitation Fellow of the Japan Society for the Promotion of Science.

## References

Ando, I., Asakawa, N. and Webb, G.A. (1998a) In *Solid State NMR of Polymers* (Eds., Ando, I. and Asakura, T.) Elsevier, Amsterdam, pp. 1–21.

Ando, I., Kameda, T., Asakawa, N., Kuroki, S. and Kurosu, H. (1998b) *J. Mol. Struct.*, **441**, 213–230.

Asakawa, N., Kurosu, H. and Ando, I. (1994a) *J. Mol. Struct.*, **323**, 279–285.

Asakawa, N., Kurosu, H., Ando, I., Shoji, A. and Ozaki, T. (1994b) *J. Mol. Struct.*, **317**, 119–129.

Beger, R.D. and Bolton, P.H. (1997) *J. Biomol. NMR*, **10**, 129–142.

Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, F., Bryce, M.D., Rogers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) *J. Mol. Biol.*, **112**, 535–542.

Blanchard, L., Hunter, C.N. and Williamson, M.P. (1997) *J. Biomol. NMR*, **9**, 389–395.

Braun, D., Wider, G. and Wüthrich, K. (1994) *J. Am. Chem. Soc.*, **116**, 8466–8469.

Clore, G.M. and Gronenborn, A.M. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 5891–5898.

de Dios, A.C. (1996) *Progr. NMR Spectrosc.*, **29**, 229–278.

de Dios, A.C., Laws, D.D. and Oldfield, E. (1994) *J. Am. Chem. Soc.*, **116**, 7784–7786.

Haigh, C.W. and Mallion, R.B. (1980) *Progr. NMR Spectrosc.*, **13**, 303–344.

Havlin, R.H., Le, H., Laws, D.D., de Dios, A.C. and Oldfield, E. (1997) *J. Am. Chem. Soc.*, **119**, 11951–11958.

Kabsch, W. and Sander, C. (1983) *Biopolymers*, **22**, 2577–2637.

Kuszewski, J., Qin, J., Gronenborn, A.M. and Clore, G.M. (1995) *J. Magn. Reson.*, **B106**, 92–96.

Kuszewski, J., Gronenborn, A.M. and Clore, G.M. (1996) *Protein Sci.*, **5**, 1067–1080.

Laws, D.D., de Dios, A.C. and Oldfield, E. (1993) *J. Biomol. NMR*, **3**, 607–612.

Le, H. and Oldfield, E. (1994) *J. Biomol. NMR*, **4**, 341–348.

Le, H., Pearson, J.G., de Dios, A.C. and Oldfield, E. (1995) *J. Am. Chem. Soc.*, **117**, 3800–3807.

Luginbühl, P., Szyperski, T. and Wüthrich, K. (1995) *J. Magn. Reson.*, **B109**, 229–233.

Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright, P.E. and Wüthrich, K. (1998) *Pure Appl. Chem.*, **70**, 117–142.
Oldfield, E. (1995) *J. Biomol. NMR*, **5**, 217–225.
Ösapay, K. and Case, D.A. (1991) *J. Am. Chem. Soc.*, **113**, 9436–9444.
Pearson, J.G., Wang, J.-F., Markley, J.L., Le, H. and Oldfield, E. (1995) *J. Am. Chem. Soc.*, **117**, 8823–8829.
Pearson, J.G., Le, H., Sanders, L.K., Godbout, N., Havlin, R.H. and Oldfield, E. (1997) *J. Am. Chem. Soc.*, **119**, 11941–11950.
Sitkoff, D. and Case, D.A. (1997) *J. Am. Chem. Soc.*, **119**, 12262–12273.
Spera, S. and Bax, A. (1991) *J. Am. Chem. Soc.*, **113**, 5490–5492.
Szilágyi, L. (1988) *Progr. NMR Spectrosc.*, **27**, 325–444.
Williamson, M.P. (1994) *Biochem. J.*, **297**, 249–260.
Williamson, M.P. and Asakura, T. (1993) *J. Magn. Reson.*, **B101**, 63–71.
Williamson, M.P. and Asakura, T. (1997) In *Methods in Molecular Biology*, Vol. 17 (Ed., Reid, D.G.) Humana Press, Totowa, NJ, pp. 53–69.
Wishart, D.S. and Sykes, B.D. (1994a) *J. Biomol. NMR*, **4**, 171–180.
Wishart, D.S. and Sykes, B.D. (1994b) *Methods Enzymol.*, **239**, 363–392.
Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S. and Sykes, B.D. (1995a) *J. Biomol. NMR*, **5**, 1–22.
Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995b) *J. Biomol. NMR*, **6**, 135–140.
Yao, J., Dyson, H.J. and Wright, P.E. (1997) *FEBS Lett.*, **419**, 285–289.