the conductance gap of individual tubes is much larger than the activation gap obtained from the resistance data. This point is under further study. Tunnelling measurements in air, rather than in vacuum, can at best be partly responsible for some of the deviations.

1. Iijima, S., *Nature*, 1991, **356**, 56–58.
2. Ebbesen, T. W. and Ajayan, P. M., *Nature*, 1992, **358**, 220–223.
3. Smalley, R. E., *Mater. Sci. Engg.*, 1993, **B19**, 1–7.
4. Tsang, S. C., Harris, P. J. F. and Green, M. L. H., *Nature*, 1993, **362**, 520–522.
5. Ajayan, P. M., Ebbesen, T. W., Ichihashi, T., Iijima, S., Tanigaki, K. and Hiura, H, *Nature*, 1993, **362**, 522–524.
6. Ajayan, P. M and Iijima, S., *Nature*, 1993, **361**, 333–334
7. Iijima, S. and Iichihashi, T., *Nature*, 1993, **363**, 603–605
8. Bethune, D. S., Kiang, C. H., de Vries, M. S., Gorman, G, Savoy, R., Vazquez, J. and Beyers, R., *Nature*, 1993, **363**, 605–607.
9. Ugarte, D., *Nature*, 1992, **359**, 707–709.
10. Mintmire, J. W., Dunlap B. I., and White, C. T., *Phys. Rev. Lett.*, 1992, **68**, 631–634.
11. Hamada, N., Sawada, S and Oshiyama, A, *Phys. Rev. Lett.*, 1992, **68**, 1579–1581.
12. Zhang, Z and Lieber, C. M., *Appl Phys. Lett*, 1993, **63**, 2792–2794.
13. Ebbesen, T W, Hiura, H., Fujita, J., Ochiai, Y., Matsui, S. and Tanigaki. K., *Chem. Phys. Lett*, 1993, **209**, 83–90.
14. Ajayan, P. M, Iichihashi, T and Iijima, S., *Chem Phys. Lett.*, 1993, **202**, 384–388
15. Oberlin, A., in *Chemistry and Physics of Carbon* (ed. Thrower, P. A), Marcel Dekker, New York, 1988, vol. 22, pp. 1–143.
16. Yoshida, M and Osawa, E, *Fullerene Sci. Technol.*, 1993, **1**, 55–74.
17. Warren, B E., *Phys. Rev.*, 1941, **59**, 693–698.
18. Ruland, W., in *Chemistry and Physics of Carbon* (ed. Walker, P. L), Marcel Dekker, New York, 1968, vol. 4, pp. 1–84.
19. Pang, L.-S. K, Saxby, J. D. and Chatfield, S. P., *J Phys. Chem.*, 1993, **97**, 6941–6942.
20. deHeer, W. A., Knight, W. D., Chou, M Y. and Cohen M. L., in *Solid State Physics* (eds. Ehrenreich, H. and Turnbull, D.), Academic Press, New York, 1987, vol. 40, pp 93–181.

# Analysis of the effects of amino acid sequence on the structure of proteins

## S. Baranidharan and M. R. N. Murthy

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

**The conformation of amino acid side chains as observed in well-determined structures of globular proteins has earlier been extensively investigated. In contrast, the structural features of the polypeptide backbone that result from the occurrence of specific amino acids along the polypeptide have not been analysed. In this article, we present the statistically significant features in the backbone geometry that appear to be a consequence of the occurrence of rotamers of different amino acid side chains by analysing 102 well-refined structures that form a random collection of proteins. It is found that the persistence of helical segments around each residue is influenced by the residue type. Several residues exert asymmetrical influence between the carboxyl and amino terminal polypeptide segments. The degree to which secondary structures depart from an average geometry also appears to depend on residue type. These departures are correlated to the corresponding Chou and Fasman parameters of amino acid residues. The frequency distribution of the side chain rotamers is influenced by polypeptide secondary structure. In turn, the rotamer conformation of side chain affects the extension of the secondary structure of the backbone. The strongest correlation is found between the occurrence of $g^+$ conformation and helix propagation on the carboxyl side of many residues.**

THE central dogma of protein folding holds that the three-dimensional structure of proteins is completely determined by their respective amino acid sequence. This implies that the occurrence of each amino acid along the polypeptide chain influences the polypeptide fold individually as well as in association with residues that occur in close proximity. In light of this, it is surprising that no detailed analyses of the stereochemical effects resulting from the occurrence of specific amino acids along the polypeptide chain are available in the literature. In contrast, the conformations of individual amino acids as observed in proteins have extensively been studied[1–6]. Janin et al.[3] analysed the high resolution protein structures available at that time and reported frequency distribution of different side chain conformations of all amino acids. It was observed that $\chi_1$ ($N - C_\alpha - C_\beta - C_\gamma$ angle) distribution is trimodal with $g^-$ conformation ($C_\gamma$ trans to $H_\alpha$) rare and $g^+$ ($C_\gamma$ trans to $C'$) position preferred in most amino acids. They also reported the estimates of deviations of the dihedral angles from the $g^-, t$ ($C_\gamma$ trans to $N$) and $g^+$ positions. Ponder and Richards[5] reexamined side chain conformations and observed that 67 distinct side chain rotamers account for most of the conformations of 17 of the 20 amino acids. Similar restricted set of conformational

states of side chains has been reported[7]. McGregor et al.[4] evaluated the frequency distribution of side chain conformations separately for helices, β-strands and non α/non β regions of proteins. They observed suppression of $g^+$ and shift towards $t$ conformation in α-helical regions when compared to non α/non β segments. This was rationalized in terms of an unacceptable contact between $C_\gamma$ and main chain carbonyl oxygen atom of one of the residues of the amino terminal helical turn, when the side chain assumes $g^-$ or $g^+$ conformation. Richardson and Richardson[8] analysed statistical preference of amino acids to occur at specific locations on α-helices and concluded that the acidic amino acids asparate and glutamate tend to occur at amino ends (NCAP residues) while the basic residues lysine, arginine and histidine tend to occur towards the carboxyl end (CCAP residues). On the other hand, the hydrophobic residues leucine, phenylalanine and methionine appear to cluster in the interior region of helices. Recently, Dunbrack and Karplus[6] have listed elaborate tables showing the distribution of rotamer conformations in different regions of Ramachandran diagram[9] as an aid to protein modeling by homology methods. Their work suggests a relationship between the conformations of the polypeptide backbone and the side chain in several 'cells' to which the Ramachandran diagram was divided for analysis. Schrauber et al.[10] examined amino acid conformations in proteins and concluded that a significant number of side chains exist in a strained state. In contrast to these studies, here we analyse the main chain conformation as viewed from a coordinate system fixed to the side chain. This approach has led to a quantitative evaluation of the impact of specific residues on secondary structure, rigidity of polypeptide fold and the effect of particular rotamers of amino acids on the stability of secondary structures. The main results of the analysis are presented in the light of earlier observations.

## Materials and methods

Structures for analysis were selected from the Brookehaven National protein data bank[11]. All structures with unique sequences solved to better than 2 Å resolution were selected for initial analysis. However some of the entries were incomplete and they were omitted. The 102 protein structures selected for final analysis are listed in Table 1.

In order to describe the main chain conformation with respect to side chain, two coordinate systems were defined, as shown in Figure 1. In both systems, the origin is at the $C_\alpha$ position of a given amino acid. System 1a (Figure 1a) allows examination of the main chain orientations seen from a system fixed with respect to $C_\alpha$, $C_\beta$ and $C_\gamma$ atoms of the side chains. Due to the

previously observed trimodal distribution of the side chain conformation in most amino acids, the main chain might be anticipated to have trimodal distribution when viewed in this system. In system 1b (Figure 1b) the polypeptide fold is viewed with respect to a coordinate system fixed to the polypeptide backbone ($N$, $C_\alpha$, $C_\beta$ atoms). The conformation of the side chain does not influence the analysis in this system. However, it will allow estimation of the rigidity of the polypeptide at each side chain position in terms of the degree of variability of the main chain fold.

For analysing the main chain conformation around a specific residue such as glutamate, the coordinates of 11 residues centred around each glutamate were transformed to the systems defined in Figure 1 such that $C_\alpha$ of all glutamates are at the origin. This transformation was performed separately for residues occurring in helices and β-strands. Identification of secondary structures was based on the method of Kabsch and Sander[12]. In this method, secondary structures are identified by the occurrence of hydrogen bonds anticipated for the respective secondary structures. The secondary structures identified by this method generally agreed with the segments reported by the original investigators. The differences were mainly confined to the terminal residues of the secondary structural segments. All the required computer programs were written in 'C' and run on a CDC machine with UNIX operating system.

Let $r_i$ be the $C_\alpha$ coordinates of $i$th residue starting from the central residue. In coordinate system 1b the quantity

$$rms_{ir} = \left( \sum_{i=1}^{N} (r_i - \langle r_i \rangle)^2 / N \right)^{1/2}. \tag{1}$$

where $N$ is the number of occurrences of the residue of interest, defines the scatter in the position of the $i$th $C_\alpha$ from its mean position. In helical segments this value should be 0 if all helices possess an ideal geometry at the residue under consideration. The actual values of $rms_{ir}$ for different residues in helices and β-strands are likely to reflect the flexibility of the secondary structure in the vicinity of the residue of interest. The quantity

$$\langle d_i \rangle = \sum_{i=1}^{N} (r_i \cdot r_i)^{1/2} / N \tag{2}$$

measures the mean distance of the $i$th residue from the residue of interest. The variation of $\langle d_i \rangle$ between different residues reflects the influence of individual residues on the polypeptide fold. An associated quantity is the rms scatter of $d_i$ defined as

$$rms_{id} = \left( \sum_{i=1}^{N} (d_i - \langle d_i \rangle)^2 / N \right)^{1/2}. \tag{3}$$

**Table 1.** List of protein data files chosen for analysis

| PDB entry | Protein name | Resolution (Å) | PDB entry | Protein name | Resolution (Å) |
|---|---|---|---|---|---|
| pdb1aap | Alzheimer's amyloid B-protein | 1.50 | pdb2had | Haloalkane dehalogenase | 1.90 |
| pdb1acx | Actinoxanthin | 2.00 | pdb2hbg | Haemoglobin (deoxy) | 1 50 |
| pdb1ak3 | Adenylate kinase isozyme-3 | 1.90 | pdb2hpr | Histidine containing | |
| pdb1alc | Alpha-lactalbumin | 1.70 | | phosphocarrier protein | 2 00 |
| pdb1ald | Aldolase-A | 2 00 | pdb2mcm | Macromycin | 1.50 |
| pdb1apb | L-arabinose binding protein | 1.76 | pdb2mhr | Myohemerythrin | 1.30 |
| pdb1bpt | Bovine pancreatic trypsin inhibitor | 1.70 | pdb2mlt | Melittin | 2.00 |
| pdb1cdp | Calcium parvalbumin | 1.60 | pdb2ovo | Ovomucoid third domain | 1.50 |
| pdb1cox | Cholesterol oxidase | 1.80 | pdb2pab | Prealbumin (human plasma) | 1.80 |
| pdb1crn | Crambin | 1.50 | pdb2pia | Phthalate deoxygenase reductase | 2.00 |
| pdb1csc | Citrate synthase | 1 70 | pdb2prk | Proteinase K | 1.50 |
| pdb1ecn | Erythrocruorin | 1.40 | pdb2rhe | Immunoglobin B-J (V-MNMR) RHE | 1 60 |
| pdb1gcr | Gamma-II crystallin | 1.60 | pdb2rsp | *Rous sarcoma* virus protease | 2 00 |
| pdb1hip | High potential iron protein | 2.00 | pdb2sar | Ribonuclease SA | 1.80 |
| pdb1hne | Elastase (human neutrophil) | 1 84 | pdb2sga | Proteinase A | 1.50 |
| pdb1hoe | Alpha-amylase inhibitor | 2.00 | pdb2sns | Staphylococcal nuclease | 1 50 |
| pdb1hyp | Hydrophobic protein from soybean | 1 80 | pdb2sod | Superoxide dismutase | 2.00 |
| pdb1ifb | Intestinal fatty acid-binding protein | 1.96 | pdb2stl | Subtilisin | 1 80 |
| pdb1mba | Myoglobin | 1.60 | pdb2tec | Thermitase/eglin-c complex | 1.98 |
| pdb1mee | Mesentericopeptidase | 2.00 | pdb2tsc | Thymidylate synthetase complex | 1 97 |
| pdb1ntp | Modified beta trypsin | 1.80 | pdb31bi | Interleukin-1 beta | 2.00 |
| pdb1nxb | Neurotoxin-B | 1.38 | pdb3app | Acid proteinase | 1 80 |
| pdb1omd | Oncomodulin (rat) | 1 85 | pdb3apr | Acid proteinase/peptide inhibitor | 1 80 |
| pdb1pal | Parvalbumin | 1 65 | pdb3b5c | Cytochrome B5 (bovine) | 1 50 |
| pdb1paz | Pseudoazurin | 1.55 | pdb3bcl | Bacteriochlorophyll A protein | 1.90 |
| pdb1pcy | Plastocyanin | 1.60 | pdb3chy | Che Y (*Salmonella typhimurium*) | 1.66 |
| pdb1pgx | Protein G-type-7 | 1.66 | pdb3cla | Chloramphenicol acetyl transferase | 1 75 |
| pdb1ppt | Avian pancreatic polypeptide | 1.37 | pdb3est | Elastase (porcine) | 1 65 |
| pdb1r69 | R1-69 *N*-terminus 434 repressors | 2.00 | pdb3fgf | Basic fibroblast growth factor | 1 60 |
| pdb1rbp | retinol binding protein (human) | 2.00 | pdb3grs | Glutathione reductase | 1.54 |
| pdb1rdg | Rubredoxin D | 1 40 | pdb3il8 | Interleukin | 2.00 |
| pdb1rei | Immunoglobin B-J V-DIMR REI | 2.00 | pdb3mcg | Immunoglobin MCG (trigonal) | 2.00 |
| pdb1rgk | Ribonuclease T1 | 1 87 | pdb3rp2 | Proteinase II (rat) | 1 90 |
| pdb1rms | Ribonuclease MS | 1.90 | pdb3tln | Thermolysin | 1 60 |
| pdb1rop | ROP: Col E1 repressor of prima | 1.70 | pdb451c | Cytochrome C551 | 1 60 |
| pdb1sn3 | Scorpion neurotoxin | 1 80 | pdb4blm | Beta-lactamase | 2.00 |
| pdb1tld | Trypsin (bovine, orthorhombic) | 1 50 | pdb4fd1 | Ferredoxin (Azotobacter. Vinel.) | 1 90 |
| pdb1ubq | Ubiquitin (human) | 1 80 | pdb4fxn | Flavodoxin (Clos. Mp. Sem | |
| pdb1utg | Uteroglobin (rabbit) | 1.34 | | Quinone) | 1 80 |
| pdb1ypi | Triose phosphate isomerase (yeast) | 1.90 | pdb4i1b | Interleukin 1B | 2 00 |
| pdb256b | Cytochrome B562 (*E coli*) | 1.40 | pdb4pep | Pepsin (porcine) | 1 80 |
| pdb2act | Actinidin | 1.70 | pdb5cyt | Cytochrome C (albacore reduced) | 1 50 |
| pdb2alp | Alpha-lytic protease | 1.70 | pdb5pti | Trypsin inhibitor | 1.00 |
| pdb2aza | Azurin | 1 80 | pdb5rxn | Rubredoxin | 1 20 |
| pdb2ccy | Cytochrome C (prime) | 1.67 | pdb5tim | Triose phosphate isomerase | 1 83 |
| pdb2cdv | Cytochrome C3 (*D. vulgaris*) | 1 80 | pdb5tnc | Troponin C (turkey) | 2 00 |
| pdb2cga | Chymotrypsinogen A | 1 80 | pdb6fab | Immunoglobin FAB | 1 90 |
| pdb2cna | Concanavalin A | 2 00 | pdb6q21 | Protein catalytic domain complex | 1 95 |
| pdb2cyp | Cytochrome C (peroxidase) | 1.70 | pdb7rsa | Ribonuclease Z (phosphate-free) | 1 26 |
| pdb2fb4 | Igg1 FAB (lambda) KOL | 1.90 | pdb8abp | ABP/D-galactose | 1 49 |
| pdb2fcr | Flavodoxin | 1.80 | pdb9pap | Papain | 1 65 |
| pdb2gbp | D-galactose-binding protein | 1 90 | pdb9wga | Wheat germ agglutinin | 1 80 |

The mean and rms deviation values of $\phi(C'_{i-1} - N_i - C_{\alpha i} - C'_i)$ and $\psi(N_i - C_{\alpha i} - C'_i - N_{i+1})$ in helical and β-strand segments for each residue type were also computed.

Glycine and alanine were not included in some of the analysis as these residues do not contain all the atoms required for transformation to system 1*b*. The quantities $\text{rms}_{tr}$, $\langle d_i \rangle$ and $\text{rms}_{id}$ were evaluated separately for residue type, residue conformation ($g^-$, $t$ and $g^+$) and secondary structure. The allocation of the side chains to $g^-$, $t$ and $g^+$ was based on $\chi_i$. Torsion angles of 0–120°, 120–240° and 240–360° were assigned to $g^-$, $t$ and $g^+$ rotamers respectively in 13 amino acids in which there is no ambiguity regarding $C_\gamma$. Instead of $C_\gamma$, S in cysteine, $C_{\gamma 1}$ in valine and isoleucine, O in serine and $O_{\gamma 1}$ in threonine were used for torsion angle calculations.
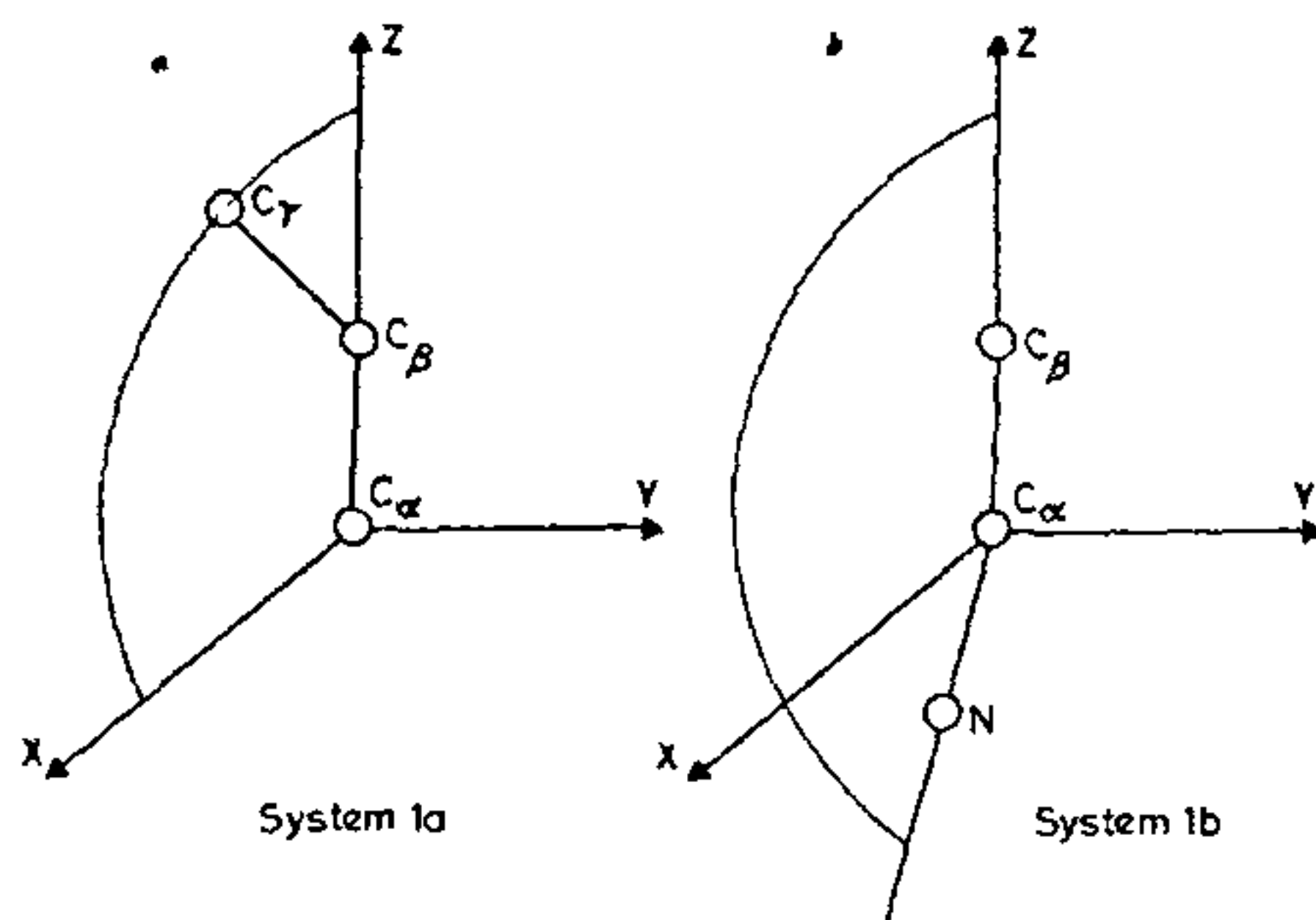
Figure 1. *a*, Cooridnate system in which $C_\alpha$ is at ongin, $C_\beta$ is along the Z-axis and $C_\gamma$ is in the X–Z plane; *b*, Coordinate system in which $C_\alpha$ is at origin, $C_\beta$ is along the Z-axis and $N$ in the X–Z plane.

During the course of this analysis, it was noticed that the tendency for secondary structures to terminate after or before specific residues depended both on residue type and residue conformation. Hence, for each amino acid occurring in helical and β-strand segments, the frequency with which the secondary structure terminates after a given number of residues was counted.

## Results and discussion

### Regularity in helices and sheets

In an ideal helix, the coordinates of the $C_\alpha$ of all residues expressed in the system 1*b* attached to a residue selected within the helical segment will be invariant. Hence the rms scatter of the neighbouring $C_\alpha$ positions ($rms_{lr}$) in this system is of interest and reflects the lack of rigidity of structure. However, a large component of the scatter could arise from errors in the coordinates and hence inclusion of only the most reliably determined structures is likely to provide useful information on flexibility. Table 2 lists the rms scatter ($rms_{lr}$) of the first neighbours in the forward (carboxyl side) and backward (amino side) directions for each of the amino acids. Also shown in Table 2 are the mean and rms values of $\phi$ and $\psi$ at each residue in helical and β-strand segments. These values were very similar for different and independent subsets of the structures chosen for analysis suggesting that the observations are not artifacts of random errors in the coordinates.

The $rms_{lr}$ scatter in the position of the neighbouring residue in the forward direction is larger than the scatter of the amino terminal neighbour in both helical and β-strand segments. The rms values are slightly larger for polar side chains. More interestingly, the observed scatter values ($rms_{lr}$) are correlated to the Chou and Fasman parameters[13] of the corresponding residues in both the helical and β-strand segments. Figures 2*a*–*d* show a plot of $rms_{lr}$ values against Chou and Fasman parameters. The correlation coefficients vary between −0.5 (sheet) and −0.8 (helix). In general, the rms values in β-strands are larger than in helical segments in conformity with the known higher flexibility of strands. The rms values of $\phi$ and $\psi$ are comparable in helical segments while $\phi$ appears to have greater flexibility than $\psi$ in β-strands.

In order to further examine the flexibility of helical segments in the context of any given residue, the mean distance $\langle d_i \rangle$ from an amino acid at position $i$ to the amino acid at position $i + 4$ (representing one full turn of the helix) and the rms scatter ($rms_{4d}$) of this distance were evaluated. The results are summarized in Table 3.

Table 3 reveals several interesting features not all of which have been analysed at present for stereochemical significance. The mean distance to the fourth neighbour in helices across all residues is 6.46 Å and is associated with an rms value of 1.1 Å. In general, it appears that the pitch of the helix is affected by the occurrence of specific residues, the larger pitch being associated with polar residues. Also alteration in the pitch is not always symmetrical. Generally, increase in pitch appears to be correlated to the degree of variation in the distance between the residues in consecutive turns in helices. Table 3 also lists the distances and rms deviations for β-strands. The mean distance and associated rms deviation of the fourth neighbour distances across all residues in β-strands are 12.95 Å and 1.1 Å respectively. In strands, $\phi$ and $\psi$ values have large variations when compared to those of helices (Table 2). However, the inter-residue distances within β-strands are not affected by these variations. In several cases, the rms deviations are smaller in β-strands when compared to helices. In contrast to α-helices, the shorter fourth neighbour distances in β-strands are associated generally with larger scatter.

### Helical persistence lengths

The number of times each residue occurs in helical segments and the number of their first and fifth neighbours that are still in the same helical segment were counted. It was observed that the frequency with which a helix continues unbroken up to a specified length (1 to 5) depended on the residue type at the origin. Table 4 lists the percentage of cases for each residue type for which the first and the fifth residues in the carboxyl and amino sides of the central residue are within the same helical segment.

It may be noticed that the number of neighbours

**Table 2.** RMS scatter (Å) from the mean position of $C_\alpha$ atoms situated next to different residue types after transformation to (main chain) coordinate system $1b$ and mean and rms deviations of $\phi$ and $\psi$ values at each residue type in helix and sheet

| Residue | Helix | | | | | | | Sheet | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | No. | $\phi$ | rms | $\psi$ | rms | amino | carbo | | No. | $\phi$ | rms | $\psi$ | rms | amino | carbo |
| Ala | 556 | −63.3 | 6 8 | −41.3 | 6.6 | − | − | | 301 | −122.1 | 28 3 | 140.7 | 19 5 | − | − |
| Arg | 175 | −63.8 | 8.0 | −42.1 | 7.9 | 0.37 | 0 63 | | 89 | −116.5 | 22.8 | 137.5 | 22 3 | 0 48 | 0.63 |
| Asn | 155 | −64.9 | 9.1 | −40.7 | 11.6 | 0.47 | 0.64 | | 78 | −108 4 | 23.2 | 132 0 | 22.0 | 0 50 | 0.73 |
| Asp | 216 | −63.7 | 6.8 | −42.5 | 10 0 | 0 35 | 0.61 | | 101 | −111.6 | 29 9 | 128 3 | 26.3 | 0 64 | 0.90 |
| Cys | 56 | −64.9 | 8 3 | −42.3 | 8 0 | 0.42 | 0 69 | | 109 | −116.7 | 23.5 | 138.7 | 26.0 | 0.55 | 0.72 |
| Gln | 192 | −63.8 | 8.0 | −41.1 | 6.9 | 0.37 | 0.55 | | 131 | −112.3 | 21.2 | 135 5 | 16.4 | 0.45 | 0.67 |
| Glu | 296 | −64.5 | 7.3 | −41 1 | 7.4 | 0.35 | 0.49 | | 124 | −110 8 | 20 7 | 133 6 | 25.0 | 0 46 | 0 66 |
| Gly | 161 | −64 3 | 7.5 | −44.8 | 19.8 | − | − | | 244 | −129.6 | 35 0 | 157.0 | 27 2 | − | − |
| His | 82 | −63.6 | 6.0 | −42 5 | 8.6 | 0 44 | 0.49 | | 48 | −113.1 | 22 1 | 139.9 | 18.0 | 0.50 | 0.73 |
| Ile | 249 | −64.3 | 8.0 | −44.0 | 6.3 | 0.34 | 0.55 | | 295 | −115.4 | 18 4 | 130.7 | 17.7 | 0.42 | 0.60 |
| Leu | 386 | −64.6 | 7.9 | −42.2 | 7.2 | 0.35 | 0.55 | | 340 | −107.5 | 21.4 | 129.8 | 18 0 | 0.50 | 0.66 |
| Lys | 317 | −64.3 | 8.7 | −42.2 | 8.3 | 0.36 | 0.50 | | 157 | −110.3 | 25.0 | 133 3 | 19.3 | 0 51 | 0.67 |
| Met | 112 | −65.2 | 8 9 | −40.7 | 7 4 | 0.30 | 0.47 | | 78 | −115.6 | 23 7 | 137 1 | 26 5 | 0.41 | 0 63 |
| Phe | 171 | −62.7 | 7.5 | −44.3 | 8 1 | 0.33 | 0.54 | | 190 | −121.7 | 24.7 | 141 1 | 20 3 | 0.50 | 0 66 |
| Pro | 83 | −59.5 | 4.2 | −35.7 | 7.9 | 0.20 | 0 59 | | 66 | − 70.4 | 8 9 | 140.9 | 21.7 | 0 24 | 0.71 |
| Ser | 250 | −65 1 | 8.8 | −40.5 | 8 8 | 0.40 | 0 82 | | 268 | −121 6 | 25.4 | 142.4 | 25.0 | 0.58 | 0.72 |
| Thr | 205 | −65.0 | 8.0 | −42.2 | 8.6 | 0 40 | 0.72 | | 292 | −117.9 | 20 8 | 138.1 | 16.9 | 0 47 | 0.56 |
| Trp | 67 | −63.1 | 7.8 | −42.7 | 7.8 | 0.30 | 0.45 | | 61 | −117.8 | 22.0 | 139.2 | 15.2 | 0.47 | 0.70 |
| Tyr | 125 | −63.4 | 8 0 | −43.6 | 9 4 | 0 45 | 0.82 | | 199 | −121.6 | 21 4 | 140 4 | 18.1 | 0.45 | 0.59 |
| Val | 276 | −64 9 | 7.0 | −43 4 | 7 2 | 0 35 | 0 65 | | 447 | −116.8 | 18.2 | 133 0 | 16 2 | 0 44 | 0 58 |

**Table 3.** Mean distance to the fourth neighbour and rms deviation of the distance for each residue type in helix and sheet

| Residue | Helix | | | | | | Sheet | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | carboxyl side | | | amino side | | | carboxyl side | | | amino side | | |
| | $N_4$ | $\langle d_4\rangle$ | rms$d_4$ | $N_4$ | $\langle d_4\rangle$ | rms$d_4$ | $N_4$ | $\langle d_4\rangle$ | rms$d_4$ | $N_4$ | $\langle d_4\rangle$ | rms$d_4$ |
| Ala | 367 | 6.25 | 0 29 | 391 | 6.19 | 0.20 | 69 | 13.25 | 0 55 | 89 | 13 23 | 0 69 |
| Arg | 112 | 6 57 | 1.26 | 138 | 6.43 | 1.08 | 28 | 13 14 | 0 46 | 22 | 12.87 | 1 42 |
| Asn | 120 | 6.71 | 1.46 | 116 | 6 65 | 1.45 | 24 | 13.12 | 0.75 | 32 | 12.58 | 1 58 |
| Asp | 173 | 6.42 | 0 99 | 122 | 6 56 | 1.39 | 32 | 12.80 | 1 58 | 47 | 12 81 | 0.88 |
| Cys | 45 | 6.57 | 1 53 | 34 | 6.45 | 0.85 | 32 | 13.16 | 0.74 | 30 | 12 86 | 1.56 |
| Gln | 143 | 6.41 | 0 95 | 122 | 6.36 | 0.93 | 40 | 12.81 | 1 25 | 34 | 13 41 | 0.38 |
| Glu | 230 | 6.30 | 0 56 | 180 | 6.31 | 0 66 | 34 | 13 01 | 1.21 | 31 | 12.45 | 1 82 |
| Gly | 124 | 6.26 | 0.20 | 99 | 6 16 | 0.43 | 63 | 12.70 | 1 58 | 60 | 13.00 | 1 11 |
| His | 55 | 6.25 | 0.30 | 60 | 6.33 | 0.60 | 19 | 12.95 | 0.78 | 19 | 12.68 | 1 43 |
| Ile | 181 | 6.32 | 0.80 | 178 | 6.51 | 1.01 | 75 | 12.95 | 0.95 | 74 | 12 89 | 1 13 |
| Leu | 224 | 6.39 | 1.15 | 301 | 6 36 | 0.87 | 72 | 12 74 | 1.29 | 102 | 12.93 | 0 90 |
| Lys | 190 | 6.40 | 1.09 | 249 | 6.33 | 0.77 | 45 | 13 00 | 0.78 | 48 | 13 18 | 0.50 |
| Met | 67 | 6.41 | 1.12 | 84 | 6 30 | 0 66 | 20 | 13.07 | 0.78 | 16 | 12 66 | 1.75 |
| Phe | 119 | 6.42 | 1.10 | 109 | 6.42 | 1 08 | 47 | 12.80 | 1 59 | 57 | 13 06 | 0 94 |
| Pro | 82 | 6 43 | 1 00 | 20 | 8 60 | 1.76 | 27 | 12.52 | 1.84 | 21 | 12.51 | 2 00 |
| Ser | 177 | 6 96 | 1 87 | 171 | 6 90 | 1.91 | 80 | 12 98 | 1 48 | 77 | 13 09 | 0 95 |
| Thr | 159 | 6 71 | 1.54 | 139 | 6.62 | 1.39 | 93 | 12 99 | 0.90 | 74 | 12 93 | 1 33 |
| Trp | 43 | 6 09 | 0.25 | 41 | 6 36 | 0 88 | 7 | 12 82 | 0 94 | 12 | 13.29 | 0 42 |
| Tyr | 92 | 6 60 | 1.41 | 94 | 7 02 | 2 06 | 59 | 13 05 | 0 83 | 43 | 12.71 | 1.78 |
| Val | 204 | 6 64 | 1.58 | 179 | 6.66 | 1.40 | 108 | 13.00 | 0.98 | 89 | 13 00 | 0 78 |

which still remain in the helical segment around certain residues is different for the forward and backward directions of the polypeptide chain. This difference may be associated with an error based on counting statistics of $(N_c + N_n)^{1/2}$ where $N_c$ and $N_n$ are the number of neighbours in the carboxyl and amino sides respectively, where the segment is still helical. Those cases in

which the quantity $|N_c - N_n|/(N_c + N_n)^{1/2}$ exceeds 1.5 are marked by asterisks.

Table 4 illustrates the opposite effects of asparate and lysine. Asparates tend to terminate helices in the backward direction while lysines break helices in the forward direction. Proline is known to occur preferably at the amino terminal region of the helices. Table 4
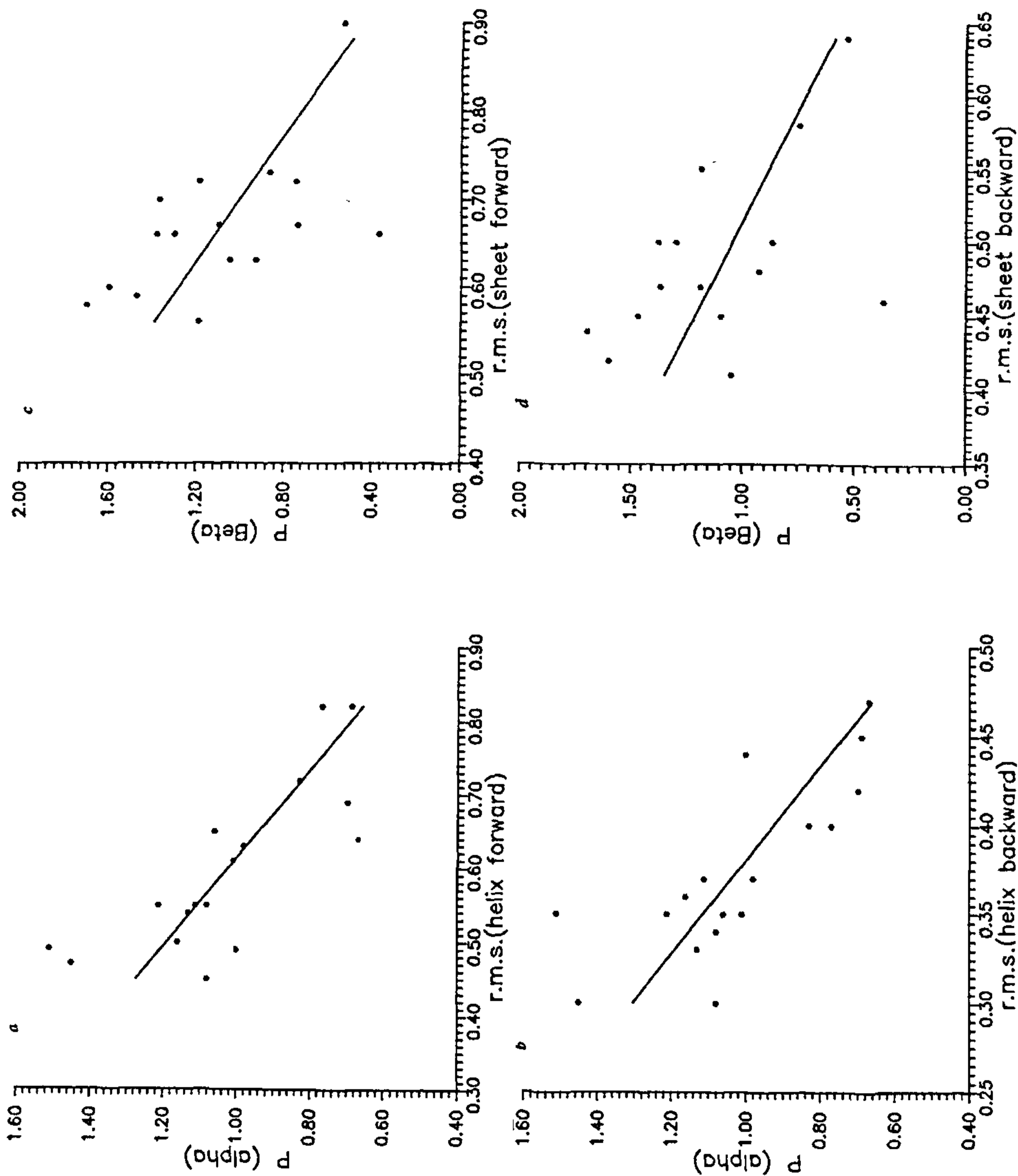
**Figure 2.** Plot showing the rms scatter in the $C_\alpha$ positions of first neighbours of amino acids (excluding glycine, alanine) occurring in helices and sheets as expressed in system 1*b*, against Chou and Fasman parameters of the respective amino acid residue. *a*, helix, carboxyl side; *b*, helix, amino side; *c*, sheet, carboxyl side; *d*, sheet, amino side.

Table 4. Percentage of instances where a helix continues unbroken after a specified number of residues from each of the amino acids. Asterisks indicate all instances where $|N_c - N_n| / (N_c + N_n)^{1/2}$ exceeds 1.5

| Residue | $N_0$ | Carboxyl side | | | | Amino side | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $N_1$ | % | $N_5$ | % | $N_1$ | % | $N_5$ | % |
| Ala | 556 | 514 | 92 | 310 | 56 | 511 | 92 | 341 | 61 |
| Arg | 175 | 163 | 93 | 100 | 57 | 171 | 98 | *122 | 70 |
| Asn | 155 | 140 | 90 | 92 | 59 | 143 | 92 | 107 | 69 |
| Asp | 216 | 200 | 93 | *154 | 71 | 191 | 88 | 109 | 50 |
| Cys | 56 | 50 | 89 | 33 | 59 | 51 | 91 | 29 | 52 |
| Gln | 192 | 178 | 93 | 114 | 59 | 176 | 92 | 103 | 54 |
| Glu | 296 | 230 | 78 | 172 | 58 | *273 | 92 | 156 | 53 |
| Gly | 161 | 155 | 96 | *115 | 71 | 140 | 87 | 89 | 55 |
| His | 82 | 70 | 85 | 45 | 55 | 77 | 94 | 55 | 67 |
| Ile | 249 | 232 | 93 | 143 | 57 | 233 | 94 | 159 | 64 |
| Leu | 386 | 355 | 92 | 196 | 51 | 364 | 94 | *265 | 69 |
| Lys | 317 | 286 | 90 | 158 | 50 | 303 | 96 | *212 | 67 |
| Met | 112 | 107 | 96 | 60 | 54 | 108 | 96 | 69 | 62 |
| Phe | 171 | 162 | 95 | 100 | 58 | 157 | 92 | 92 | 54 |
| Pro | 83 | *83 | 100 | *75 | 90 | 41 | 49 | 19 | 23 |
| Ser | 250 | 225 | 90 | 167 | 67 | 228 | 91 | 152 | 61 |
| Thr | 205 | 180 | 88 | *130 | 63 | 172 | 84 | 100 | 49 |
| Trp | 67 | 64 | 96 | 38 | 57 | 63 | 94 | 36 | 54 |
| Tyr | 125 | 112 | 90 | 77 | 62 | 110 | 88 | 93 | 74 |
| Val | 276 | 264 | 96 | *186 | 67 | 258 | 93 | 157 | 57 |

further shows that the helix remains unbroken in the carboxyl side more frequently when a proline occurs in the helical segment compared to the occurrence of any other residue. About 90% of the fifth neighbours of proline is still in helical segments while this number varies from 50% (lysine) to 71% (asparate) for other residues. In contrast, proline is the strongest helix breaker in the backward direction. Threonine and valine occur more frequently at the amino terminal region of helical segments.

These observations are in general agreement with those of Richardson and Richardson[8]. These investigators classified different amino acids based on their preference to occur at different positions of helical segments. Asparagine, asparate and serine were found to occur preferentially at amino terminal positions and hence' were designated as NCAP residues. In contrast, glutamate, lysine and histidine were found frequently as penultimate residues of helices (CCAP-1 position). Glycine occurred at both NCAP and CCAP positions. However, Table 4 shows that glycine prefers NCAP position. This apparent discrepancy in the frequency of occurrence of glycine at termini arises due to the difference in the definition of helical segments in the present analysis and that of Richardson and Richardson[8]. Many of the glycines included as carboxyl terminal residues in their analysis are excluded from helical segments·on hydrogen bonding criterion used here. In the present analysis lysine exhibits a strong tendency to terminate helices on the carboxyl side. In the earlier analysis, although lysine is not found as a CCAP residue, it occurred

more frequently at CCAP-1, CCAP-2 and CCAP-3 positions. Similar preferences were observed in the earlier investigations on the occurrence of glutamate and arginine. However, for these residues, in terms of the continuation of the helices up to 5 residues, the differences in the forward and backward directions are detectable but are not very significant. Similarly, serine was observed as an NCAP residue. However, it occurs frequently at CCAP, CCAP-1, CCAP-2 positions such that in terms of helix persistence propensity, no significant difference exists in the forward and backward directions. The additional strong tendency observed here is that of leucine which prefers carboxyl ends and valine and threonine which prefer amino ends.

## Polypeptide in the context of side chain rotamer

It is well known that most of the side chains can assume three major rotamer conformations with respect to $\chi_1$. More recently Dunbrack and Karplus[6] have observed a strong correlation between side chain rotamer conformation and backbone geometry. In order to examine the effect of the choice of a particular rotamer on the main chain conformation, the coordinates of the polypeptide were transformed to the coordinate system 1a. The resulting coordinates of 11 residues centred around each residue type were plotted. These plots clearly revealed trimodal distribution of the main chain propagation anticipated due to the preferred conformations of the side chain[4]. However, secondary structure of the main

chain appeared to break off in certain directions more frequently than in other directions suggesting a relationship between side chain conformation and secondary structure stability. Table 5 lists the occurrences of each residue in $g^+$, $t$ and $g^-$ conformation in helical segments along with the number of cases where the secondary structure persists after 5 residues towards the carboxyl and amino ends from the central residue.

The preferences of each amino acid for $g^+$, $t$ and $g^-$ rotamers as deduced from Table 5 are in good agreement with the observations of McGregor et al.[4] and Dunbrack and Karplus[6]. However, due to the larger database used in this analysis compared to McGregor et al.[4] and smaller number of cells to which rotamers are assigned compared to those of Dunbrack and Karplus[6] the statistical preferences of different conformations appear to be more informative in the present case. As observed in the earlier analysis, $g^-$ conformation is strongly suppressed in helical regions except for residues serine, threonine and proline (not analysed by McGregor et al.[4]). Compared to non-helical regions (data not shown), the preference for $t$ conformation is enhanced in helical segments.

Table 5 shows the strong dependence of helix propagation on the rotamer conformation of leucine residues. On the carboxyl side of leucine, the helix terminates within 5 residues in 60% of the cases in $g^+$ conformation compared to 32% of the cases in $t$ conformation. Hence, apart from the shift towards $t$ conformation in helical segments, the stability of the helix itself appears to be reduced due to the occurrence of $g^+$ conformation. This effect is also discernible in arginine, histidine and isoleucine. Threonine occurs predominantly in the $g^+$ con-

formation in helices. The reduced occurrence of $g^-$ conformation is also accompanied by more frequent disruption of the secondary structure on the carboxyl side. In contrast to the significant correlation between residue rotamer conformation and helix propagation on the carboxyl side, the amino side of the residues appears to be insensitive. In leucine and isoleucine, there is a weak tendency for helix termination on the amino side when the residues are in $t$ conformation rather than the $g^+$ conformation. In this context it might be of interest to examine the side chain rotamer frequencies as a function of the residue position in the helix. However, with the size of the data base available at present, such an analysis may not be statistically significant.

In order to understand the stereochemical origin of the above observations, ideal helices (poly-alanine) of length 11 residues were built using the Insight-II software. The residue of interest was located at the sixth position in the desired conformation. It was observed that leucine side chain when present in the $g^+$ conformation makes severe van der Waals contacts with atoms of the neighbouring amino terminal residues for most values of $\chi_2$ (Figure 3). However, no such severe contacts are made with the residues on the amino or carboxyl side when the residue is in $t$ conformation. This would suggest premature termination of the helix on the amino terminal side of leucine when it occurs in $g^+$ conformation. However, it is not consistent with the statistical trends revealed by Table 5 where the occurrence of $g^+$ conformation disrupts the helix propagation on the carboxyl side. Hence, the stereochemical or other reasons that lead to some of the features observed in Table 5 are not clear.

Table 5. Occurrence of side chain rotamers in helical segments and the number of first and fifth neighbours in the same helical segment

| Residue | $N_0$ | | | Carboxyl side | | | | | | Amino side | | | | | |
| | | | | $N_1$ | | | $N_5$ | | | $N_1$ | | | $N_5$ | | |
| | $g^-$ | $t$ | $g^+$ | $g^-$ | $t$ | $g^+$ | $g^-$ | $t$ | $g^+$ | $g^-$ | $t$ | $g^+$ | $g^-$ | $t$ | $g^+$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Arg | 9 | 82 | 84 | 9 | 78 | 76 | 7 | 54 | 39 | 7 | 81 | 83 | 5 | 56 | 61 |
| Asn | 4 | 21 | 130 | 4 | 20 | 116 | 3 | 12 | 77 | 2 | 21 | 120 | 2 | 18 | 87 |
| Asp | 15 | 40 | 161 | 15 | 35 | 150 | 14 | 24 | 116 | 12 | 35 | 144 | 2 | 22 | 85 |
| Cys | 2 | 22 | 32 | 2 | 22 | 26 | 2 | 13 | 18 | 0 | 22 | 29 | 0 | 11 | 18 |
| Gln | 4 | 86 | 102 | 4 | 82 | 92 | 4 | 50 | 60 | 2 | 79 | 95 | 2 | 50 | 51 |
| Glu | 14 | 108 | 174 | 13 | 102 | 115 | 10 | 71 | 101 | 11 | 97 | 165 | 4 | 60 | 92 |
| His | 3 | 34 | 45 | 3 | 33 | 34 | 3 | 22 | 20 | 1 | 32 | 44 | 0 | 23 | 32 |
| Ile | 15 | 15 | 219 | 8 | 15 | 209 | 5 | 15 | 123 | 15 | 8 | 210 | 14 | 4 | 141 |
| Leu | 1 | 140 | 245 | 1 | 136 | 218 | 1 | 96 | 99 | 1 | 126 | 237 | 1 | 87 | 177 |
| Lys | 13 | 142 | 162 | 11 | 134 | 141 | 5 | 74 | 79 | 13 | 131 | 159 | 9 | 91 | 112 |
| Met | 1 | 38 | 73 | 0 | 36 | 71 | 0 | 16 | 44 | 1 | 36 | 71 | 1 | 24 | 44 |
| Phe | 5 | 104 | 62 | 5 | 102 | 55 | 5 | 62 | 33 | 2 | 96 | 59 | 1 | 52 | 39 |
| Pro | 26 | 0 | 57 | 26 | 0 | 57 | 26 | 0 | 49 | 14 | 0 | 37 | 7 | 0 | 12 |
| Ser | 92 | 60 | 98 | 79 | 60 | 86 | 64 | 40 | 63 | 82 | 54 | 92 | 56 | 42 | 54 |
| Thr | 48 | 7 | 150 | 34 | 5 | 141 | 21 | 5 | 104 | 34 | 6 | 132 | 30 | 1 | 69 |
| Trp | 7 | 40 | 20 | 7 | 39 | 18 | 7 | 21 | 10 | 6 | 38 | 19 | 0 | 25 | 11 |
| Tyr | 3 | 70 | 52 | 3 | 69 | 40 | 3 | 43 | 31 | 2 | 66 | 42 | 2 | 55 | 36 |
| Val | 12 | 233 | 31 | 10 | 227 | 27 | 9 | 156 | 21 | 7 | 222 | 29 | 6 | 132 | 19 |
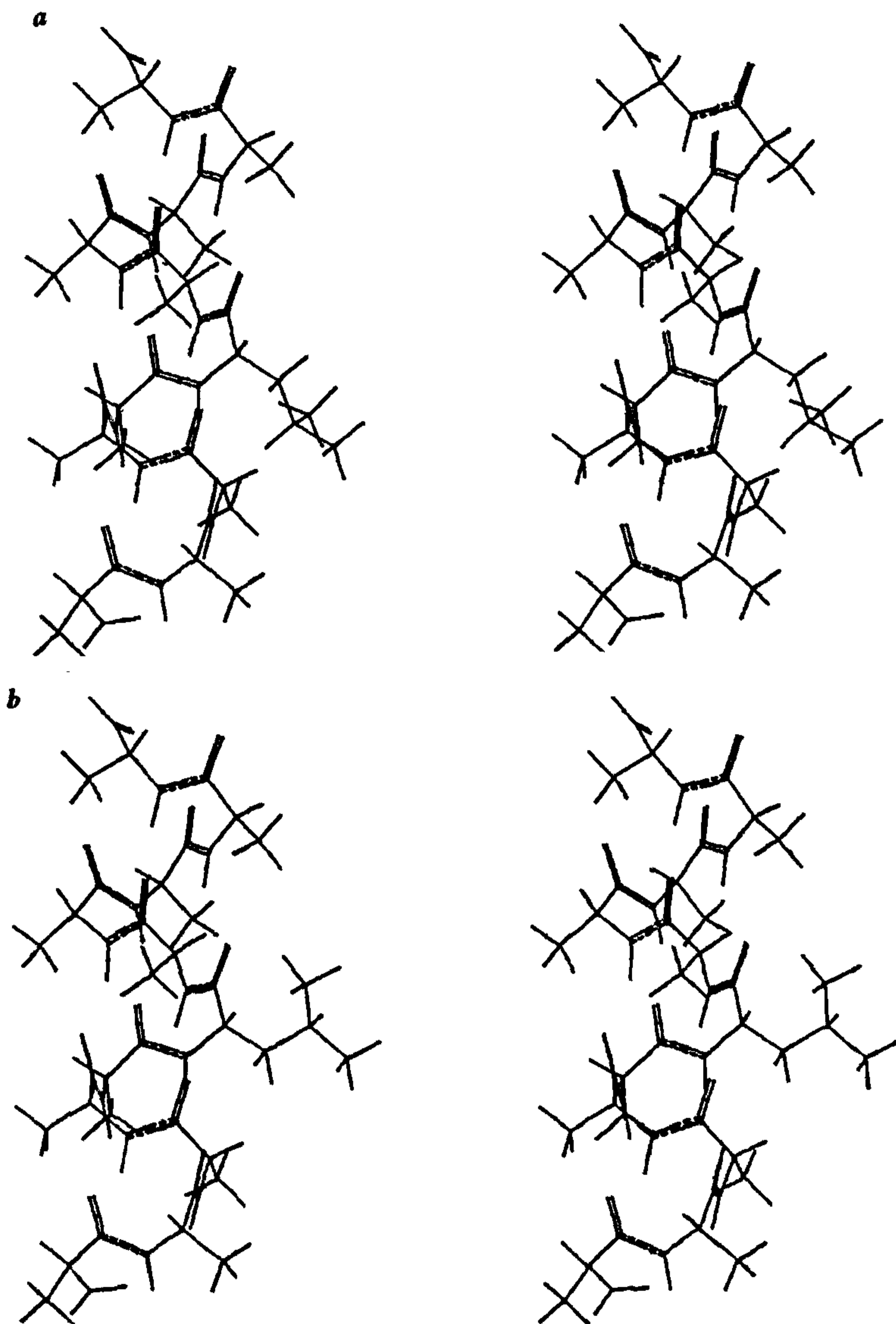
Figure 3. Two rotamers of leucine in an ideal helix. *a*, The side chain of leucine (sixth residue in a hypothetical ideal helix) makes short contacts with the residues on the amino terminal side when $\chi_1$ is in $g^+$ conformation. The torsion angle $\chi_2$ is shown in *t* conformation; *b*, The side chain in *t* conformation does not interfere with the residues in the helix.

## Conclusions

In this article we have presented an analysis of 102 well-determined structures with a view to understanding the possible consequences of the occurrence of specific amino acids along the polypeptide chain. Several statistically significant and interesting features have emerged from these studies. The analysis of helical persistence lengths around each residue type is generally in good agreement with earlier studies by Richardson and Richardson[8] on the preference of amino acids for positions on helical segments. The present analysis provides some additional features and insights. The preference of certain amino acids to occur at termini of helical segments could be understood in terms of the asymmetrical impact of each side chain on the polypeptide backbone. The departure of helical and β-strands from an average geometry is dependent on the residue type. This departure is correlated to the Chou and Fasman parameters[13] of the amino acid for the respective secondary structure. The conformation of the amino acid side chain, which is probably strongly influenced by packing interactions in the folded structure, appears to influence the stability of the secondary structure. In particular, occurrence of $g^+$ rotamers influences the helix propagation on the carboxyl side of certain residues. However, individual side chains do not appear to have remarkably significant and predictable effect on the backbone geometry in the case of helices and sheets. Analysis of the effect of rotamer conformation on polypeptide backbone in non α/non β region of the polypeptide might reveal more definitive correlations. The results presented here provide a starting point for further analysis of these effects in terms of specific interactions of the amino acid side chains with different conformation of the polypeptide backbone.

1. Ponnuswamy, P. K. and Sasisekharan, V., *Biopolymers*, 1971, **10**, 565–582.
2. Sasisekharan, V. and Ponnuswamy, P. K., *Biopolymers*, 1971, **10**, 583–592.
3. Janin, J., Wodak, S., Levitt, M. and Maigret, B., *J. Mol. Biol.*, 1978, **125**, 357–386.
4. McGregor, M. J., Islam, S. A., Sternberg, M. J. E., *J. Mol. Biol.*, 1987, **198**, 295–310.
5. Ponder, J. W. and Richards, F. M., *J. Mol. Biol.*, 1987, **193**, 775–791.
6. Dunbrack, R. L. and Karplus, M., *J. Mol. Biol.*, 1993, **230**, 543–574.
7. Tuffery, P., Etchebest, C., Hazout, S. and Lavery, R., *J. Biomol. Struct. Dyn.*, 1991, **8**, 1267–1289
8. Richardson, J. S. and Richardson, D. C., *Science*, 1988, **240**, 1648–1652.
9. Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V., *J. Mol. Biol.*, 1963, **7**, 95–99.
10. Schrauber, H., Eisenhaber, F. and Argos, P., *J. Mol. Biol.*, 1993, **230**, 592–612.
11. Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., Weng, J. C., *Protein Data Bank in Crystallographic Databases: Information Content, Software Systems, Scientific Applications* (eds. Allen, F. H., Bergerhoft, G. and Sievers, R.), Bonn, Chester, Cambridge; Int. Union of Crystallography, 1987, pp. 107–132.
12. Kabsch, W. and Sander, C., *Biopolymers*, 1983, **22**, 2577–2637.
13. Chou, P. Y. and Fasman, G. D., *Biochemistry*, 1974, **13**, 211–222.

# Carbon and oxygen isotopic compositions of the regionally metamorphosed Archaean carbonate rocks of the Dharwar craton: A preliminary appraisal

S. Das Sharma, R. Srinivasan, S. M. Ahmad and D. J. Patil

National Geophysical Research Institute, Uppal Road, Hyderabad 500 007, India.

Carbon and oxygen isotopic data of the carbonate rocks from the greenschist to upper amphibolite facies metamorphosed Archaean Dharwar Supergroup show that, while the carbon isotopic compositions have remained unaltered, there is a 5‰ depletion in $\delta^{18}O$ ratios in response to increase in the grade of metamorphism. Exchange with metamorphic pore water at temperatures in excess of 400°C is considered to have caused this shift.

THE primary carbon and oxygen isotope ratios are prone to alteration by post-depositional diagenetic and metamorphic processes. It is therefore necessary to screen the data for secondary alteration that may have affected the primary signatures before using isotopic ratios to infer biogeochemical processes and environments prevalent at the time of deposition. With this aim we have undertaken a study of the carbon and oxygen isotopic compositions in the carbonate rocks of different metamorphic grades of the Archaean supracrustal belts of the Dharwar craton.

The > 2.5 Ga Archaean supracrustal rocks of the Dharwar Supergroup[1] have been divided into the lower Bababudan and upper Chitradurga Groups. The Bababudan Group comprises metamorphosed quartz arenite–basalt–rhyolite–BIF association. This is unconformably overlain by the Chitradurga Group which has been subdivided into Vanivilas, Ingaldhal and Hiriyur formations. The Vanivilas Formation constituting the lower part of the Chitradurga Group is composed of quartz arenite–carbonate association with minor mafic sills intruding the sequence. These are overlain by pillow basalts, rhyolites and tuffs of the Ingaldhal Formation and greywackes of the Hiriyur Formation.

Approximately E–W trending isograds transect the NNW–SSE trending Dharwar supracrustal belts, such that greenschist facies rocks are exposed in the north, amphibolite in the middle and granulite facies rocks in the south[2,3] (Figure 1). In the amphibolite/granulite grade terrain, the supracrustal rocks have been considered by some workers[4,5] to belong to a sequence older than the Dharwar Supergroup—namely the Sargur Group. However, Srinivasan[6] and Naha et al.[7] have provided evidence to show that the supracrustal rocks of the Sargur Group
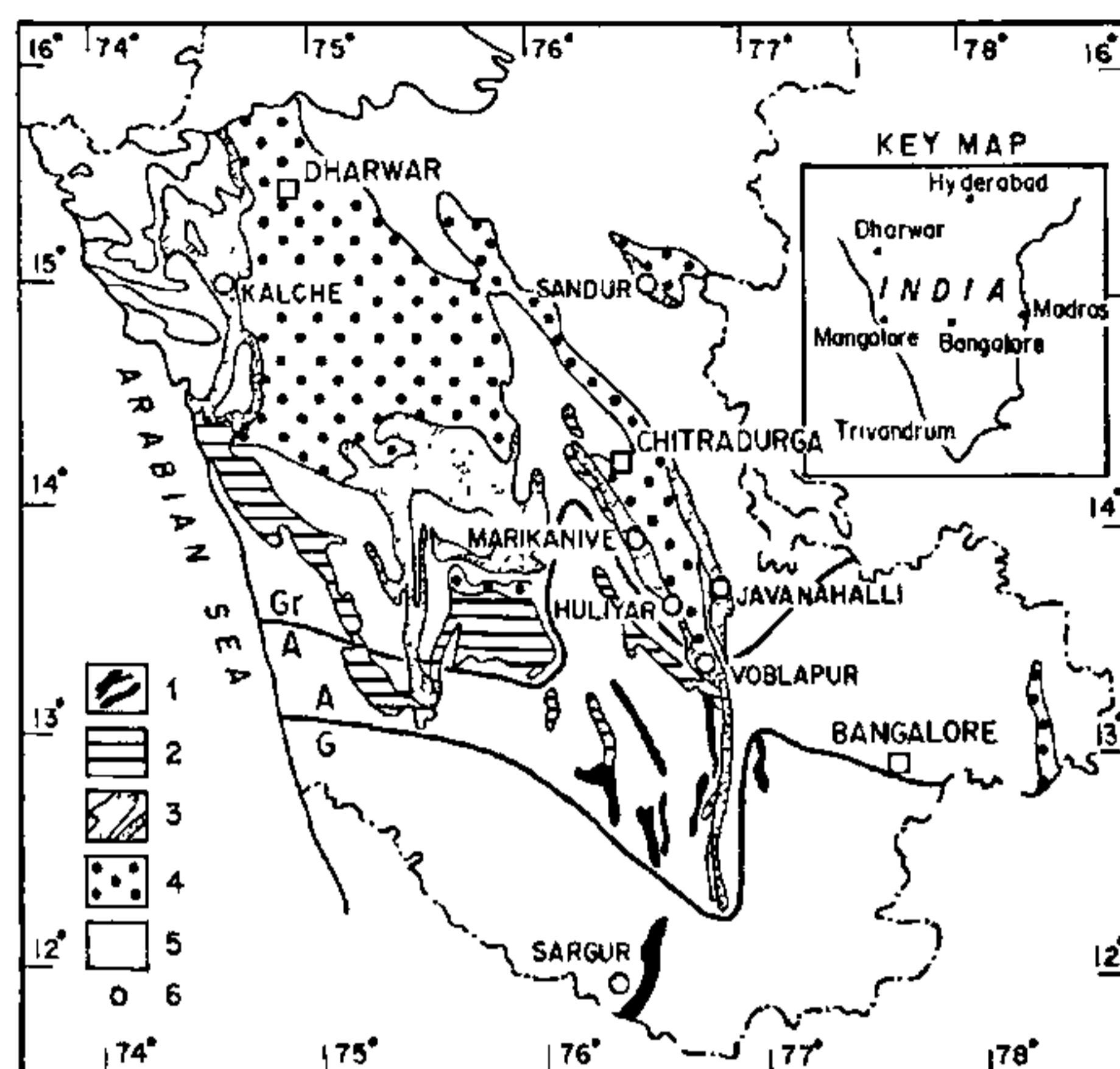


Figure 1. Geological map of the Dharwar craton showing sample locations 1, Sargur supracrustal belts; 2, Bababudan Group of Dharwar Supergroup, 3, Vanivilas Formation of Chitradurga Group of Dharwar Supergroup; 4, Ingaldhal volcanic and Hinyur Formations of Dharwar Supergroup; 5, Undifferentiated gneisses, granites and granulites; 6, Sample locations. Heavy lines Gr/A and A/G are metamorphic isograds between greenschist-amphibolite and amphibolite-granulite respectively (after Raase et al.[3])

represent higher grade metamorphic equivalents of those of the Dharwar Supergroup, which view has been adopted in this paper. The classification of the supracrustals is summarized in Table 1. The carbonate rocks under study belong to the Vanivilas Formation in the greenschist to low grade amphibolite facies metamorphosed Dharwar supracrustal belts (low grade metamorphic terrain) and also by their stratigraphic equivalents in the high grade Sargur supracrustal belts (high grade metamorphic terrain).

Limestones are generally cement grey to greyish white in colour in the greenschist facies terrain. They are microcrystalline to coarse grained and are composed of calcite, dolomite, minor amounts of epidote, phlogopite, zoisite, quartz and rarely graphite. By contrast, in the amphibolite and granulite facies these are coarsely crystalline marbles usually greyish white to white in colour. They are composed of calcite, dolomite, ankerite, tremolitic amphibole, garnet, phlogopite and rarely forsterite. Dolomites of low grade metamorphic terrain are usually cherty in which beds or laminae of chert alternate with dolomitic layers. Dolomite as well as chert layers are usually microcrystalline. Despite deformation and low grade metamorphism that has affected these carbonate rocks, at places they show well-preserved domical as well as columnar forms of stromatolites[8]. Small quartz veins are seen traversing the cherty dolomite. Discrete metapelitic interlayers with minor graphite are sometimes

857