Int. J. Peptide Protein Res. 19, 1982, 1–9

Side chain characteristic main chain conformations of amino acid residues

A.S. KOLASKAR and V. RAMABRAHMAM

School of Life Sciences, University of Hyderabad, Hyderabad, India

Received 23 April, accepted for publication 17 July 1981

Main chain conformation characteristics of the respective side chains of the 20 naturally occurring amino acid residues were obtained by the analysis of (ϕ, ψ)-data from the crystal structures of 38 different globular proteins. The following observations are of interest: (i) For amino acids other than Glu, Thr and aliphatic amino acids, at least one main chain conformation is stabilised solely by side chain atoms. Such conformations are listed. (ii) In globular proteins, the main chain conformations which are significantly destabilised by side chain atom interactions are observed. The stabilising force for those conformations seems to come from main chain atom interactions. (iii) A set of conformations for which the main chain and side chain interactions are almost equal but in opposite directions, is listed and these conformations will be taken by residues mainly due to the effect of surroundings.

The results can be used to study the folding of polypeptide chains and they also provide an insight into the role of the respective side chains on main chain conformations of amino acid residues.

Key words: data analysis; protein crystal structure; side chain specificity of amino acid residues.

In an attempt to understand the folding of polypeptides and proteins, we have analysed the crystal structure data of a large number of different globular proteins. It is our belief that the wealth of information hidden in these crystal structure data can be fruitfully used in predicting the folding of polypeptide chains. Several groups (1–7) are also concentrating their efforts in this direction and have derived very useful information.

In our first attempt (8) we derived the torsional potential energy functions using these data, which, when used along with other accepted potential energy functions, such as non-bonded and electrostatic, have reproduced (ϕ, ψ)-dipeptide maps which are very similar to the empirical map obtained by Pohl (9). This has prompted us to analyse the (ϕ, ψ)-probability maps obtained for individual amino acid residues. The comparisons of these probability maps have indicated that the main chain conformational similarity is as much an independent property of amino acid residues as their secondary structure forming capacities (10). We observed that the (ϕ, ψ)-probability distributions are unique for Gly and Pro. And in other cases, though the side chains are different, their effect on the

0367–8377/82/010001–09 $02.00/0

respective (ϕ, ψ)-distributions seems to be less than the effect of side chains of Gly and Pro.

There is thus a need to get information on the conformations in (ϕ, ψ)-plane which are significantly affected by the group of atoms which form the respective side chains. This can be done in two ways. One is by use of semi-empirical potential energy functions to calculate the dipeptide maps, as has been done by several groups (11–14) and to deduct the contribution of the potential energy due to the main chain atoms. Thus, one can arrive at a set of conformations which are more or less solely stabilised due to the presence of a particular side chain. This approach has its limitations since the potential functions used by different groups are different and the accuracy of them is doubtful (15).

The other approach, which we have followed, is purely empirical, using the available (ϕ, ψ)-data from crystal structures of different globular proteins. In order to find those conformations in the (ϕ, ψ)-plane that are significantly affected by respective side chains of amino acid residues, via the above-mentioned approach, we have developed an algorithm based on the following logic. The (ϕ, ψ)-distribution obtained for each residue from a large number of globular proteins having different three-dimensional structures and functions represents the conformational property of that amino acid residue. The environment of a given residue taking up a particular conformation being different in different proteins, the inter-residue interactions are masked in this (ϕ, ψ)-distribution. Thus, in short, though the folding process of any polypeptide chain may be assumed to be guided mainly by intra-residue and inter-residue interactions, the (ϕ, ψ)-probability maps can be used to study the intra-residue interactions, which are due to the main chain and side chain atoms of a given residue. In the present analysis, the interactions from main chain atoms have been deduced from the (ϕ, ψ)-probability map of each residue and the conformations which are affected only by side chain atoms are derived. The conformations obtained in this manner help us understand the role played by each of these amino acid residues during secondary and tertiary structure formations.

"GENERAL-L-RESIDUE" PROBABILITY MAP

ψ \ φ	-180°	-160°	-140°	-120°	-100°	-80°	-60°	-40°	-20°	0°	20°	40°	60°	80°	100°	120°	140°	160°	180°
180°	0.4	1.1	1.4	1.1	1.0	0.6	0.6	.03	.04	0	.02	.01	.08	.05	.06	.05	0.1	0.2
160°	0.6	1.4	2.5	2.7	1.8	1.7	0.9	0.2	.02	.03	.01	.01	.06	.05	.06	.02	.08	0.2
140°	0.4	1.3	2.8	3.2	2.4	1.8	1.4	0.2	0.1	0.3	0	0	.02	.09	.05	.08	.05	0.1
120°	0.2	0.5	1.0	1.9	1.2	1.0	0.5	0.2	.03	.06	.04	.02	.04	.05	.03	0	.05	0.1
100°	.08	0.2	0.6	0.7	0.6	0.5	0.2	.07	.05	.07	.02	.07	.06	.05	.03	.05	.01	.01
80°	.08	0.2	0.3	0.3	0.3	0.1	.05	.03	.06	0.1	.04	0.2	0.1	.07	.04	.07	.07	0
60°	.04	0.2	0.3	0.3	0.2	.06	0.1	0	.01	.09	0.1	0.3	0.2	0.1	.04	.05	.03	.09
40°	.07	0.2	0.2	0.6	0.5	0.2	.08	0	0	.01	.01	0.2	0.4	0.2	.07	.05	.02	.01
20°	.04	0.2	0.5	0.8	1.1	0.5	0.2	.02	.03	.02	.02	0.2	0.2	0.3	.07	.02	.03	0
0°	0	0.2	0.5	1.0	1.9	2.2	0.8	.08	.01	.01	0	.05	0.1	0.1	0.1	.02	.02	.03
-20°	.07	0.1	0.2	0.8	1.8	4.2	2.9	0.4	.04	0	0	0	.03	0.1	0.2	.04	.06	.01
-40°	.02	0.1	0.2	0.3	1.2	4.0	6.3	1.4	0.2	.03	0	.02	.02	.08	.04	.07	.01	.04
-60°	.04	0.1	.08	0.2	0.3	0.9	2.7	1.1	0.2	.09	0.1	.07	.08	.07	.02	.04	.02	.04
-80°	.01	.03	.04	0.1	.04	0.1	0.3	0.2	0.2	.06	.03	.03	.02	.03	.03	.04	.01	.07
-100°	.03	.02	.05	.09	.05	.09	0.1	.05	.08	0	0.1	.05	.03	.06	.05	.05	0	.03
-120°	.04	.05	.09	0.1	.09	.02	.02	.01	.04	0	.04	.03	.04	.05	.04	.03	.02	.02
-140°	.04	0.1	0.2	0.1	0.2	.09	.02	.03	.01	0	0	.03	.08	.02	.04	.03	.03	.06
-160°	0.2	0.5	0.4	0.7	0.3	0.2	0.1	.02	.01	.01	.03	.02	.03	.05	.01	.04	0.1	0.1
-180°																			

FIGURE 1

Normalised (ϕ, ψ)-probability map obtained from crystal structure data of 38 globular proteins at a grid interval of 20°. Note that the individual normalised (ϕ, ψ)-probability maps of each of the 20 residues were used. The regions A and B comprise ϕ-variations between − 140° and 0°, − 180° and 0° respectively and ψ-variations between − 100° and 0° and 80° and 180° respectively. These regions are marked in the figure.

METHODS

The (ϕ, ψ)-values of amino acid residues were collected from the crystal structure data of 38 different globular proteins. These (ϕ, ψ)-values were supplied on micro-fiche by Richard Feldmann (16). The proteins considered, together with the resolution of their crystal structures, are given in our earlier paper (10). The observed (ϕ, ψ)-probability maps were prepared individually for the 20 amino acid residues at a grid interval of 20°. This interval was chosen since the (ϕ, ψ)-values obtained from crystal structure data are accurate only up to 20° for most of the cases. The observed probability map of each residue was normalised, P_j^i representing the normalised probability of jth grid in (ϕ, ψ)-plane of the ith residue. Thus, the index j varies over all the 324 grid points in the (ϕ, ψ)-map of each residue. Using these P_j^i values, the probability P_j^{GR} is obtained

using the relation:

$$P_j^{GR} = \frac{\sum_{i=1}^{20} P_j^i}{20} \quad (1)$$

The (ϕ, ψ)-map, thus obtained, is hereafter referred to as "general L-residue" (ϕ, ψ)-probability map and is given in Fig. 1. This "general L-residue" is made up of normal main chain atoms and a fictitious side chain. This side chain is the resultant effect of all 20 side chains getting 0.05 weightage from each individual side chain. In other words, this "general L-residue" (ϕ, ψ)-probability map represents main chain atom interactions with unit weightage, while the interactions among side chain atoms themselves or interactions among side chain-main chain atoms of any individual residue with very low weightage. The normalised (ϕ, ψ)-probability maps of individual residues, excepting that of Gly, were compared grid-wise with the "general L-residue" map. The comparison with Gly (ϕ, ψ)-map was done after suitably altering the "general L-residue" map, since Gly (ϕ, ψ)-probability map possesses an inversion symmetry. The "difference probability" is represented by:

$$\Delta P_j^i = P_j^i - P_j^{GR} \quad (2)$$

These "difference probability" maps were computed for each residue. Standard deviation of ΔP_j^i was obtained for each residue using the following relation.

$$\overline{\Delta P^i} = \sum_{j=1}^{324} \Delta P_j^i / 324 \quad (3a)$$

$$\sigma_i = \sum_{j=1}^{324} (\Delta P_j^i - \overline{\Delta P^i})/323 \quad (3b)$$

Each ΔP_j^i was compared with $2\sigma_i$ and if $|\Delta P_j^i| > 2\sigma_i$, then only P_j^i value is considered to be significantly different from P_j^{GR} value and the corresponding (ϕ_j, ψ_j)-grid was assumed to have a significant contribution from the side chain of the residue and these conformations are referred as "side chain characteristic" conformations.

RESULTS AND DISCUSSION

As mentioned ΔP_j^i values are indicators of contribution to intra-residue interactions due to the side chain atoms of a particular residue, as they are obtained by taking out the effect of other influences on main chain conformations in the form of "general L-residue" probability. However, conformations having $|\Delta P_j^i| > 2\sigma_i$ can only be considered with 95% confidence limit as the conformations having major contribution from side chain atoms. These conformations for each of the 20 amino acid residues, along with corresponding ΔP_j^i values, are listed in Table 1. Table 1 shows that only a small portion of the allowed conformational space is significantly affected by side chain atoms and this effect is different for different side chains, which can be intuitively perceived. It can be seen from Table 1 that ΔP_j^i values are not always positive. $\Delta P_j^i > 0$ indicates that the corresponding main chain conformation is stabilised by interactions due to respective side chain atoms. Similarly, the respective side chain atoms destabilise the main chain conformations if $\Delta P_j^i < 0$. Thus, the data given in Table 1 can be categorised into four types, namely:

(a) Side chain and main chain stabilised conformations

$$P_j^i > 2\sigma_i; \quad P_j^{GR} > 2\sigma_i; \quad \Delta P_j^i > 0$$

(b) Side chain stabilised – main chain indifferent conformations

$$P_j^i > 2\sigma_i; \quad P_j^{GR} < 2\sigma_i; \quad \Delta P_j^i > 0$$

(c) Side chain destabilised – main chain stabilised conformations

$$P_j^i > 2\sigma_i; \quad P_j^{GR} > 2\sigma_i; \quad \Delta P_j^i < 0$$

(d) Side chain destabilised – main chain indifferent conformations

$$P_j^i < 2\sigma_i; \quad P_j^{GR} > 2\sigma_i; \quad \Delta P_j^i < 0$$

Conformations of each residue falling under each of these categories are listed respectively in the four rows (a) to (d) of Table 1. These are further illustrated below.

Case (a) (side and main stabilised)

$$P_j^i > 2\sigma_i; \quad P_j^{GR} > 2\sigma_i; \quad \Delta P_j^i > 0$$

TABLE 1

"Side chain characteristic" conformations from protein crystal structure data. These conformations are divided into four categories (for details see text) and the corresponding ΔP_j^i values are given beneath. $2\sigma_i$ (σ_i = standard deviation) is also given together with residue. All grid values are in degrees

Ala $2\sigma_i = 0.76$

(a) Side and main stabilized

(−80, −60)	(−60, −60)	(−40, −60)	(−80, −40)
1.78	3.49	1.43	2.00
(−60, −40)	(−80, −20)	(−160, 160)	
2.07	0.89	1.05	

(b) Side stabilised – main indifferent ——

(c) Side destabilised – main stabilised

(−140, 120)	(−120, 120)	(−140, 140)
−1.38	−1.44	−1.38

(d) Side destabilised – main indifferent

(−120, 100)	(−140, 160)
−1.6	−1.09

Arg $2\sigma_i = 0.74$

(a) Side and main stabilised

(−80, −60)	(−80, −20)	(−100, 120)	(−120, 140)
3.01	0.77	2.11	1.29
(−140, 160)	(−60, 160)		
1.10	0.95		

(b) Side stabilised – main indifferent

(20, −80)	(−140, −20)	(−160, 60)
0.90	1.04	0.76

(c) Side destabilised – main stabilised

(−60, −40)	(−140, 120)	(−120, 120)	(−100, 140)
−0.92	−1.78	−1.65	−0.81

(d) Side destabilised – main indifferent

(−100, 0)	(−60, 120)	(−80, 140)
−1.06	−1.38	−1.2

Asn $2\sigma_i = 0.92$

(a) Side and main stabilised ——

(b) Side stabilised – main indifferent

(−140, −160)	(−120, 0)	(80, 0)	(−120, 20)
1.04	1.29	1.24	1.24
(−100, 20)	(40, 20)	(40, 40)	(−140, 80)
1.88	1.56	2.04	1.17

(c) Side destabilised – main stabilised

(−60, −80)	(−80, −60)	(−60, −60)	(−80, −40)
−1.48	−1.90	−4.24	−1.55
(−140, 120)	(−120, 140)		
−1.29	−1.52		

(d) Side destabilised – main indifferent

(−160, 140)	(−80, 140)	(−100, 160)
−1.39	−1.11	−1.02

Asp $2\sigma_i = 0.64$

(a) Side and main stabilised

(−40, −80)	(−80, −60)	(−120, −20)	(−100, −20)
0.89	0.89	0.70	1.96

(b) Side stabilised – main indifferent

(−160, 0)	(−140, 160)	(−100, 80)	(−80, 80)
0.81	0.69	0.66	1.01

(c) Side destabilised – main stabilised

(−60, −60)	(−120, 100)	(−140, 120)	(−120, 120)
−0.72	−0.69	−1.31	−0.71
(−80, 120)	(−140, 140)	(−120, 140)	(−100, 140)
−1.07	−1.74	−0.75	−1.08
(−80, 140)			
−0.73			

(d) Side destabilised – main indifferent

(−160, 120)	(−160, 140)	(−140, 160)	(−120, 160)
−1.04	−0.91	−0.91	−0.89

Cys $2\sigma_i = 0.84$

(a) Side and main stabilised

(−120, −20)	(−90, −20)	(−60, 120)	(−80, 120)
1.29	1.20	1.59	1.65
(−160, 140)	(−140, 140)	(−120, 140)	(−100, 140)
2.63	0.98	1.89	2.79
(−140, 160)			
2.30			

(b) Side stabilised – main indifferent

(−160, −180)	(−120, −60)
1.7	1.38

(c) Side destabilised – main stabilised

(−60, −80)	(−80, −60)	(−60, −60)	(−120, 120)
−1.52	−1.11	−1.15	−0.85

(d) Side destabilised – main indifferent

(−80, −80)	(−100, 100)	(−80, 140)
−0.86	−1.23	−1.13

Glu $2\sigma_i = 0.82$

(a) Side and main stabilised

(−80, −80)	(−40, −80)	(−100, −60)	(−80, −60)
0.85	2.63	1.10	2.83
(−60, −60)	(−40, −60)	(−100, −40)	(−80, −40)
1.91	0.9	1.03	2.31
(−100, −20)	(−60, −20)	(−80, 140)	
1.18	0.95	1.99	

(b) Side stabilised – main indifferent ——

(c) Side destabilised – main stabilised

(−140, 120)	(−100, 120)	(−140, 140)	(−120, 140)
−1.07	−1.54	−1.05	−1.29

(d) Side destabilised – main indifferent

(−100, 100)	(−160, 140)	(−140, 160)
−0.95	−1.39	−0.83

Gln $2\sigma_i = 0.68$

(a) Side and main stabilised

(−60, −80)	(−40, −80)	(−100, −40)	(−80, −40)
0.94	1.10	1.44	1.19
(−120, −20)	(−140, 100)	(−140, 120)	(−100, 120)
0.79	1.16	0.83	1.58
(−60, 120)	(−140, 140)	(−100, 140)	
0.79	1.14	1.44	

(b) Side stabilised – main indifferent

(−140, 0)
1.34

(c) Side destabilised – main stabilised

(−80, −60)	(−120, 100)	(−120, 140)
−1.10	−1.19	−1.26

(d) Side destabilised – main indifferent

(−120, −40)	(−100, 0)	(−160, 140)
−0.77	−1.06	−1.03

Gly $2\sigma_i = 1.3$

(a) Side and main stabilised ——

(b) Side stabilised – main indifferent

(−100, −180)	(100, −40)	(60, −20)	(100, −20)
1.64	1.60	1.47	1.58
(60, 0)	(80, 0)	(100, 0)	(60, 20)
2.20	2.20	1.46	2.76

(c) Side destabilised – main stabilised ——

(d) Side destabilised – main indifferent

(−120, 160)	(80, 40)	(60, 60)	(80, 60)
−1.50	−1.70	−2.70	−1.91
(−140, 120)	(−120, 120)		
−1.50	−1.55		

His $2\sigma_i = 0.86$

(a) Side and main stabilised

(−80, −60)	(−60, −60)	(−100, −20)	(−100, 0)
1.86	1.28	2.15	2.45
(−80, 100)	(−160, 140)	(−80, 140)	(−120, 160)
1.89	1.53	1.22	1.21

(b) Side stabilised – main indifferent

(−120, −180)	(−80, −160)	(−160, 0)	(100, 20)
1.06	1.08	1.01	1.10
(−120, 40)			
0.91			

(c) Side stabilised – main stabilised

(−140, 120)	(−120, 120)	(−100, 120)	(−120, 140)
−1.02	−1.98	−1.22	−1.54

(d) Side destabilised – main indifferent

(−100, −60)	(−100, −40)	(−120, 100)	(−160, 120)
−1.17	−1.22	−1.33	−1.28
(−60, 140)			
−0.87			

Ile $2\sigma_i = 0.88$

(a) Side and main stabilised

(−60, −60)	(−120, 100)	(−160, 120)	(−140, 120)
1.81	2.55	1.34	4.31
(−120, 120)	(−140, 140)		
4.2	0.94		

(b) Side stabilised – main indifferent ——

(c) Side destabilised – main stabilised

(−100, −20)
−0.89

(d) Side destabilised – main indifferent

(−80, 140)
−1.44

Leu $2\sigma_i = 0.64$

(a) Side and main stabilised

(−80, −80)	(−100, −60)	(−60, −60)	(−80, −40)
0.97	1.21	1.92	0.72
(−120, 100)	(−100, 100)	(−160, 120)	(−120, 120)
1.20	1.70	0.73	0.88
(−100, 120)	(−100, 140)	(−120, 160)	
2.19	1.30	0.70	

(b) Side stabilised – main indifferent ——

(c) Side destabilised – main stabilised

(−60, −40)	(−120, 140)	(−140, 160)
−0.90	−1.06	−0.67

(d) Side destabilised – main indifferent

(−120, −180)	(−120, −20)
−0.70	−0.83

Lys $2\sigma_i = 0.52$

(a) Side and main stabilised

(−120, −180)	(−60, −80)	(−40, −80)	(−80, −60)
0.58	1.35	0.76	0.58
(−40, −60)	(−60, −40)	(−100, 0)	(−140, 160)
0.63	1.65	0.58	2.20

(b) Side stabilised – main indifferent

(−20, −60)	(−100, 60)	(−180, 100)	(−40, 100)
1.04	0.58	0.73	0.75

(c) Side destabilised – main stabilised

(−120, 100)	(−60, 120)	(−160, 140)	(−140, 140)
−0.63	−0.65	−0.66	−1.55
(−120, 160)			
−0.58			

(d) Side destabilised – main indifferent

(−60, 100)
−0.56

Met $2\sigma_i = 1.32$

(a) Side and main stabilised

(−60, −60)	(−40, −60)	(−140, 140)	(−120, 140)
9.06	1.51	2.34	3.06
(−140, 160)			
1.49			

(b) Side stabilised – main indifferent

(−100, 80)
1.36

(c) Side destabilised – main stabilised ——

(d) Side destabilised – main indifferent

(−100, −20)	(−100, 140)
−1.94	−1.81

Phe $2\sigma_i = 0.58$

(a) Side and main stabilised

(−60, −80)	(−100, −60)	(−40, −60)	(−120, −40)
0.87	0.80	0.59	0.80
(−120, −20)	(−120, 100)	(−180, 140)	(−160, 160)
1.74	1.24	0.59	1.62
(−120, 160)			
1.23			

(b) Side stabilised – main indifferent

(−80, −100)	(−180, 60)	(−180, 120)
0.66	0.71	0.76

(c) Side destabilised – main stabilised

(−60, −60)	(−100, −20)
−0.81	−1.15

(d) Side destabilised – main indifferent

(−40, −80)	(−60, −20)	(−100, 160)
−1.07	−0.76	−0.63

Pro $2\sigma_i = 1.76$

(a) Side and main stabilised

(−60, −80)	(−60, −40)
2.79	2.87

(b) Side stabilised – main indifferent

(−80, −180)	(−60, −20)	(−60, 100)	(−60, 120)
2.01	4.71	2.65	5.37

(−80, 140)	(−60, 140)	(−80, 160)	(−60, 160)
6.01	5.56	2.89	4.91

(c) Side destabilised – main stabilised ——
(d) Side destabilised – main indifferent

(−80, −60)	(−140, 120)	(−120, 120)	(−140, 140)
−2.7	−2.78	−3.15	−2.15
(−120, 140)			
−2.38			

Ser $2\sigma_i = 0.58$

(a) Side and main stabilised

(−120, 0)	(−160, 140)	(−140, 140)	(−120, 160)
1.16	1.46	1.74	0.67
(−100, 160)			
0.93			

(b) Side stabilised – main indifferent

(−160, −180)			
0.70			

(c) Side destabilised – main stabilised

(−60, −80)	(−60, −60)	(−80, −40)	(−120, 100)
−1.77	−1.97	−0.62	−1.31
(−120, 120)	(−80, 120)		
−1.50	−0.60		

(d) Side destabilised – main indifferent

(−60, 120)	(−120, 80)		
−0.93	−0.68		

Thr $2\sigma_i = 0.66$

(a) Side and main stabilised

(−120, −40)	(−100, −20)	(−160, 120)	(−140, 120)
0.87	1.76	0.77	0.92
(−120, 120)	(−120, 140)	(−120, 160)	(−100, 160)
2.39	1.40	1.13	1.24
(−80, 160)			
0.79			

(b) Side stabilised – main indifferent ——
(c) Side destabilised – main stabilised

(−80, −60)	(−60, −60)	(−60, −40)	
−2.34	−1.60	−1.07	

(d) Side destabilised – main indifferent

(−100, −60)	(−40, −60)		
−0.76	−0.76		

Trp $2\sigma_i = 0.92$

(a) Side and main stabilised

(−160, −180)	(−100, −40)	(−60, −40)	(−80, 100)
1.7	1.13	1.49	1.17
(−160, 120)	(−140, 120)	(−120, 120)	
1.66	2.37	2.73	

(b) Side stabilised – main indifferent

(−140, −180)	(−160, 20)	(−120, 20)	(−100, 20)
1.78	1.26	1.66	0.97
(−140, 80)	(−160, 100)	(−180, 160)	
1.59	0.97	1.05	

(c) Side destabilised – main stabilised

(−60, −60)			
−2.65			

(d) Side destabilised – main indifferent

(−40, −60)	(−100, −20)	(−100, 100)	(−80, 120)
−1.38	−1.21	−1.23	−1.06
(−160, 160)	(−100, 160)		
−1.13	−1.02		

Tyr $2\sigma_i = 0.90$

(a) Side and main stabilised

(−120, 100)	(−140, 120)	(−120, 120)	(−180, 140)
1.69	1.90	1.52	1.93
(−140, 140)	(−120, 140)	(−160, 160)	
1.13	3.41	1.74	

(b) Side stabilised – main indifferent

(−120, −140)	(−140, −20)	(−120, 80)	(−80, 80)
0.97	0.98	1.11	0.98

(c) Side destabilised – main stabilised

(−60, −60)	(−80, −40)	(−80, −20)	
−2.01	−2.07	−1.15	

(d) Side destabilised – main indifferent

(−40, −60)	(−100, −40)	(−160, 120)	(−80, 140)
−1.02	−1.09	−0.92	−0.98

Val $2\sigma_i = 0.86$

(a) Side and main stabilised

(−100, −60)	(−120, 80)	(−140, 100)	(−120, 100)
0.97	0.96	1.79	3.52
(−140, 120)	(−120, 120)	(−100, 120)	(−160, 140)
1.50	3.26	1.23	1.07
(−140, 140)	(−120, 140)	(−140, 160)	
1.97	1.73	1.89	

(b) Side stabilised – main indifferent ——
(c) Side destabilised – main stabilised

(−60, −60)	(−80, −20)		
−1.06	−1.08		

(d) Side destabilised – main indifferent

(−100, −40)	(−100, −20)	(−100, 0)	(−80, 140)
−1.15	−1.12	−1.06	−1.05

Conformations listed in row (a) of Table 1 are stabilised by interactions from main chain atoms ($P_j^{GR} > 0$) and side chain atoms ($\Delta P_j^i > 0$). The global energy minimum conformation for the dipeptide map for a residue having type (a) conformations is expected to be one among these conformations. For example, the observed probability maps of Ala and Ile residues show that the global energy minimum for these residues are respectively at (− 60°, − 60°) and (− 120°, 120°) grids.

Table 1 also shows that residues Asn and Gly do not possess type (a) conformation. This means that the main chain and side chain atom interactions are never in phase for these two residues when both the interactions are contributing significantly.

The analysis of ΔP_j^i values for conformations listed in row (a) for residues which prefer secondary structures such as α-helix or β-sheet indicates certain interesting features. For example, the ΔP_j^i values are maximum in α-

helical region for the α-helix preferrers, Ala, Glu, Met, Gln, His, Leu and Lys for the conformations in the grids (−60°, −60°), (−80°, −60°), (−60°, −60°), (−100°, −40°) (−100°, −20°), (−60°, −60°) and (−60°, −60°) respectively. This indicates that although for most of these residues the contribution for stabilisation from the respective side chains is significant for the α-helix conformation, namely the (−60°, −60°) grid, it is not maximum for this conformation for all the preferrers. Thus, this analysis helps explain the spread of conformations characteristic of the side chains, in this α-helical region occurring either in the same protein or different proteins. This point is further illustrated by considering a specific case of Glu residue occurring in the α-helix. It is found that Glu occurs 152 times in the α-helix in the sequences of the 38 proteins considered. Out of 152 times, it occurred 18 times in the (−80°, −60°) grid, 25 times in the (−60°, −60°) grid, 17 times in the (−80°, −40°) grid and a similarly considerable number of times in other grids of the α-helical region.

Observation of conformations occurring in row (a) reveals that although residues are not preferrers of a particular secondary structure, their side chain and main chain atom interactions can be in phase for conformations which lie in that secondary structural region of (ϕ, ψ)-plane. For example Trp and Thr are known respectively as indifferent to and a breaker of the α-helix (17). However, the (−60°, −40°) conformational grid of Trp and the (−100°, −20°) grid of Thr are stabilised by side chain as well as main chain atoms.

Case (b) (side stabilised – main indifferent)

$P_j^i > 2\sigma_i$; $P_j^{GR} < 2\sigma_i$; $\Delta P_j^i > 0$

These conformations mentioned in row (b) of Table 1 have $P_j^{GR} \approx 0$, indicating that the contribution from main chain atoms for them is negligible and the stabilisation is solely due to specific side chain. Thus these conformations are of the type which distinguish one residue from the other.

Pro is found (10) to have distinct (ϕ, ψ)-probability distribution when compared to other residues. The side chain of Pro being rigid and cyclic in nature, it will have significant effect on the respective main chain conformations. Conformations in the grids (−80°, −180°), (−60°, −20°), (−60°, 100°), (−60°, 120°), (−80°, 140°), (−60°, 140°), (−80°, 160°) and (−60°, 160°) are examples. Thus, the conformations in the range of ϕ from −40° to −80° and ψ almost over the whole range of allowed conformations are stabilised by the side chain of Pro.

There is another set of residues like Ala, Glu, Ile, Leu, Thr and Val, which do not have a single conformation in row (b), indicating that the side chains of these residues do not stabilise any conformation solely, and stabilisation from main chain atoms is essential for conformations to be taken up by them in polypeptides.

Case (c) (side destabilised – main stabilised)

$P_j^i > 2\sigma_i$; $P_j^{GR} > 2\sigma_i$; $\Delta P_j^i < 0$

The conformations which fall under this category have $P_j^{GR} > P_j^i$. Here the main chain and side chain atom interactions are out of phase. Thus, although the contribution from side chain atoms is significant in this case, it is the main chain atom interactions that make residues take up these conformations.

The case for Cys taking up the (−60°, −60°) conformation is shown by the fact that though the side chain of Cys does not prefer this conformation ($\Delta P_j^i = -1.2$), there are a considerable number of Cys residues in proteins adopting the conformation and they occur also in regular structures such as the α-helix. In fact, analysis of the data of proteins considered here shows that Cys has adopted the (−60°, −60°) conformation five times while occurring 31 times in the α-helix. Residues Gly, Met and Pro do not have a single conformation under this category.

Case (d) (side destabilised – main indifferent)

$P_j^i < 2\sigma_i$; $P_j^{GR} > 2\sigma_i$; $\Delta P_j^i < 0$

For conformations of this type the stabilising main chain interaction term gets almost nullified by the side chain destabilising interactions. A conformation of this type, taken up by a

particular residue in a polypeptide chain results from a delicate balance between main chain and side chain atom interactions, acting in opposite directions, and the effect of the neighbouring residues. Thus the inter-residue interactions may play quite a significant role when such conformations are taken up by the amino acid residue. We are investigating the neighbourhood of amino acid residues adopting this type of conformations.

Conformations in the allowed regions of the (ϕ, ψ)-plane that do not figure in Table 1 for respective residues may be assumed to be neither stabilised nor destabilised, and hence get neutral influence from side chain atoms. It is obviously the main chain atom interactions that stabilise these conformations. Thus, these conformations are not characteristic of any particular side chain.

Table 1 further shows that there are a few conformations which can be either stabilised or destabilised by several side chains. For example, the grids ($-100°$, $-20°$), ($-80°$, $-60°$), ($-60°$, $-60°$) which lie in the region A of the (ϕ, ψ)-map and ($-140°$, $140°$), ($-120°$, $120°$), ($-140°$, $120°$), ($-120°$, $100°$), ($-120°$, $140°$) and ($-80°$, $140°$) which lie in the region B, receive a significant contribution from the side chains of 10 or more residues. All these grids not only are part of thickly populated regions of the (ϕ, ψ)-plane but they are also representatives of the two major secondary structures, α-helix and β-sheet, of globular proteins. It is interesting to note that the conformations mentioned above, which are part of region A, where the main chain atoms are densely packed, show a significant effect from the side chain atoms of the residue. In addition, the α-helix preferrers get their stabilisation from side chain atom interactions for at least one of these conformations, the exception being that of Cys. Also the residues which are categorised either as breakers of or indifferent to the α-helix (17) need not necessarily be the ones getting the destabilisation effect from side chains for these conformations. In case of conformations occurring in the β-sheet, the reasons for the significant contribution from the side chain of residues to these conformations for which the main chains have extended conformation is difficult to explain. Therefore, these conformations require a more detailed study. However, a point to be noted here is that for the residues which are β-sheet preferrers, their respective side chains prefer one or the other of the conformations cited above.

A careful observation of Table 1 reveals a few conformations which are specific to a side chain. These are shown in Fig. 2. Most of the side chain specific conformations are those which have a stabilising effect on the particular conformation. Fig. 2 also points out that only 12 amino acid residues possess side chain specific conformations. In addition to Gly, residues Asn, Arg and His also have the conformation specific to their side chains in the right part of (ϕ, ψ)-map. Those residues which do not have specific conformations include the aliphatic amino acids and the hydroxy amino acid residues Thr and Ser.

The discussion of the results mentioned above brings out that the "side chain charac-

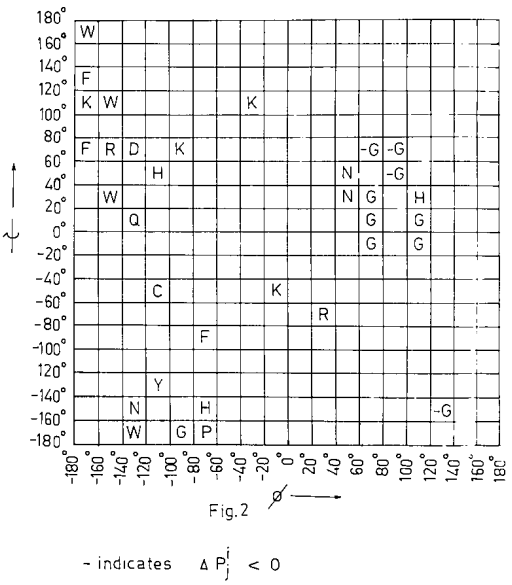

FIGURE 2

Residue mentioned in a grid of (ϕ, ψ)-map takes that conformation which has a significant contribution from its side chain. Side chain of any other residue does not contribute significantly for this grid. — sign before a residue indicates that this conformation is destabilised by respective side chain. Single letter amino acid code has been used.

teristic conformations" for each residue are the ones which should be used in studies on the folding of proteins. Further it points out that the atoms beyond C^{β} of the side chain also have a significant effect on the main chain conformation, in contrast to the previous results obtained by semi-empirical and quantum chemical calculations on dipeptides (18, 19). Thus, for the first time, conformations specific to certain amino acid residues have been obtained. This study divides some of the observed conformations in (ϕ, ψ)-plane into two distinct categories, namely those that are affected by the presence of a particular side chain and a few others that are stabilised solely by side chain interactions. In addition to this a set of conformations for each residue are listed which receive significant contributions from side chain atoms and are stabilised by main chain atom interactions. The results thus help understanding to a certain extent of the role played by side chain atoms in the tertiary structure formation of polypeptides and proteins.

REFERENCES

1. Rossmann, M.G., Liljas, A., Branden, C.I. & Banaszak, L.J. (1975) in *The Enzymes* (Boyer, P.D., ed.), 3rd edn., vol. 11, pp. 61–102, Academic Press, New York
2. Levitt, M. & Chothia, C. (1976) *Nature* **261**, 552–557
3. Richardson, J.S. (1976) *Proc. Natl. Acad. Sci. US* **73**, 2619–2623
4. Janin, J. (1979) *Nature* **277**, 491–492
5. Schulz, G.E. & Schirmer, R.H. (1978) *Principles of Protein Structure,* Springer-Verlag, New York
6. Prabhakaran, M. & Ponnuswamy, P.K. (1979) *J. Theor. Biol.* **80**, 445–450
7. Janin, J. & Chothia, C. (1980) *J. Mol. Biol.* **143**, 95–128
8. Kolaskar, A.S. & Prashanth, D. (1979) *Int. J. Peptide Protein Res.* **14**, 81–87
9. Pohl, F.M. (1971) *Nature New Biol.* **234**, 277–279
10. Kolaskar, A.S. & Ramabrahmam, V. (1981) *Int. J. Biol. Macromol.* **3**, 171–178
11. De Santis, P., Giglio, E., Liquori, A.M. & Ripamonti, A. (1965) *Nature* **206**, 456–458
12. Flory, P.J. (1969) *Statistical Mechanics of Chain Molecules,* Wiley Interscience, New York
13. Ramachandran, G.N. & Sasisekharan, V. (1968) *Advan. Protein Chem.* **23**, 283–437
14. Scheraga, H.A. (1968) *Adv. Phys. Org. Chem.* **6**, 103–184
15. Ramachandran, G.N. (1973) in *Conformation of Biological Molecules and Polymers* (Bergmann, E.D. & Pullman, B., eds.), pp. 1–13, Academic Press, Jerusalem
16. Feldmann, R. (1976) *Atlas of Macromolecular Structure on Microfiche,* 1st edn. Tracer Jitco Inc., Rockville, MD, 20852
17. Levitt, M. (1978) *Biochemistry* **17**, 4277–4285
18. Ponnuswamy, P.K. & Sasisekharan, V. (1971) *Biopolymers* **10**, 565
19. Pullman, B. & Pullman, A. (1974) *Adv. Protein Chem.* **28**, 347–526

Address:

Dr. *A.S. Kolaskar*
School of Life Sciences
University of Hyderabad
Hyderabad – 500 134
India