

# Energy Minimization Studies on $\alpha$ -Turns

C. RAMAKRISHNAN and D. V. NATARAJ

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

Received 24 March 1997

Accepted 22 July 1997

**Abstract:** Using a grid search technique, the entire conformational space of a system of four linked peptide units (tetrapeptide) was scanned to pick out geometrically possible 5→1 type hydrogen-bonded conformations defined as an  $\alpha$ -turn. The energy minimization of these conformations led to 23 distinct minimum energy conformations (MECs) falling in 13 different classes. The presence of  $\beta$  and  $\gamma$  turn type hydrogen bonds along with 5→1 type hydrogen bond gave conformational variability in a given class. The occurrence of bifurcated hydrogen bonding network was a characteristic feature of most of the MECs. In many prototype MECs non-glycyl residues such as Ala and Pro could be accommodated. Comparison of MECs with the  $\alpha$ -turn examples that are observed in proteins showed that the conformationally worked out MECs occurred in isolation in proteins, with the  $\alpha$ -helical  $\alpha$ -turn being distinctly the most predominant. © 1998 European Peptide Society and John Wiley & Sons, Ltd.

**Keywords:** Hydrogen-bonded-tetrapeptide;  $\alpha$ -turns;  $\alpha$ -turn conformation; energy minimization; secondary structure

## INTRODUCTION

Many globular proteins contain the well-known  $\alpha$ -helix, which is a succession of 5→1 hydrogen-bonded tetrapeptide segments. In a recent communication from this group [1], it has been shown that such 5→1 hydrogen-bonded segments, known as  $\alpha$ -turns, can occur in isolation. If the conformations at the  $\alpha$ -carbon atoms are close to those in  $\alpha$ -helix, they are described as ' $\alpha$ -helical  $\alpha$ -turns'. It has also been shown [1] that there are other families of  $\alpha$ -turns with non-helical conformations (though with fewer examples), occurring in globular proteins. This has prompted us to work out the possible low-energy conformations of  $\alpha$ -turns. This study is similar to those in which theoretical models were obtained by Venkatachalam and others as early as in 1968 [2, 3] for classical  $\beta$ -turns. In this paper, the methodology and results of energy minimization studies carried out on  $\alpha$ -turns are presented.

## METHODS

The principal aim of this study is to work out conformations of all possible combinations of the main torsion angles ( $\phi$ ,  $\psi$ ) at the three intervening  $\alpha$ -carbon atoms, which will result in the formation of 5→1 bond in a system of all *trans* tetrapeptide segment and at the same time will be free of any backbone steric hindrance. Hence glycyl residues are chosen at all the three positions, as the side chain imposes least conformational restriction. (Cf. the next section for the effect of other residues such as Ala and Pro). Initially a system of four linked peptide units is generated with ideal bond lengths and bond angles of a *trans* peptide unit [4] using a locally developed Fortran program. A schematic diagram of an  $\alpha$ -turn is shown in Figure 1, from which it can be seen that the specification of an  $\alpha$ -turn requires six variable parameters, viz. three sets of torsion angles, on each at  $C_2^\alpha$ ,  $C_3^\alpha$  and  $C_4^\alpha$ . By systematically varying each of these six torsion angles ( $\phi_2$ ,  $\psi_2$ ), ( $\phi_3$ ,  $\psi_3$ ) and ( $\phi_4$ ,  $\psi_4$ ) from  $-180^\circ$  to  $+150^\circ$  at intervals of  $30^\circ$ , it is possible to generate  $12^6$  combinatorial possibilities in the six-dimensional conformational hyperspace. The criteria used

Correspondence to: C. Ramakrishnan, Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

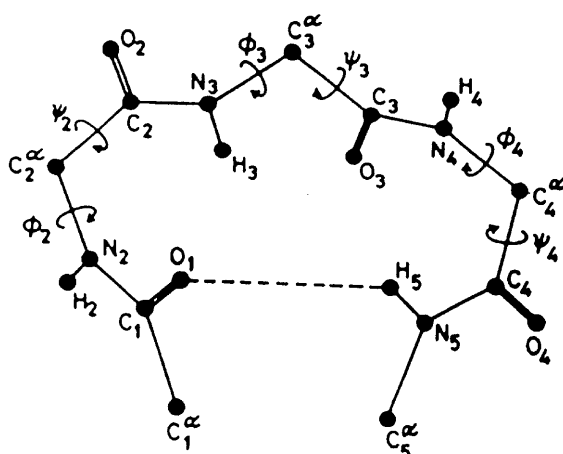


Figure 1 Schematic diagram of an  $\alpha$ -turn. The backbone atoms and torsion angles are marked.

for finding out the presence of hydrogen bonds are (i) the N...O distance must lie between 2.5 Å and 3.2 Å and (ii) the angle H-N...O must be  $\leq 30^\circ$ . When the conformations so generated are checked for the presence of 5 $\rightarrow$ 1 hydrogen bond, it turns out that only about 4000 conformations possess the requisite 5 $\rightarrow$ 1 hydrogen bond.

Although these conformations possess the necessary hydrogen bond, some may not be energetically favourable owing to steric hindrance. In order to pick up only those conformers that have a relatively low energy, they were subjected to energy minimization. To find an energy minimum in a force field it is necessary to have a proper set of potential functions and a proper iterative algorithm.

The total potential energy of a system is the sum of the contributions due to (i) non-bonded and electrostatic interactions, (ii) the distortion potential pertaining to bond lengths and bond angles, (iii) the torsional potential arising from the barrier to internal rotation about various bonds and (iv) the potential energy due to the hydrogen bond. The non-bonded potential function,  $V_{nb}$ , used in the minimization is a Buckingham six-exponential type, of the form,

$$V_{nb} = -A/r^6 + Be^{-\mu r}$$

and the electrostatic potential,  $V_{es}$ , of the form

$$V_{es} = q_i q_j / \epsilon r$$

where  $q_i$ ,  $q_j$  are the partial charges on the atoms and  $r$  is the interatomic distance. A value of 4 has been used for the dielectric constant  $\epsilon$ . The constants for

the various potential functions as well as the form and the constants for the hydrogen bond potential function are the same as those given in Ramachandran and Sasisekharan (1968) [5] and have been used extensively in earlier studies from our group [6].

The iterative algorithm used in energy minimization is a combination of steepest descent and Davidon-Fletcher-Powell (DFP) methods [7-9]. Minimization was carried out in two steps:

1. Initial minimization of geometrically generated conformations through 100 cycles of steepest descent.
2. Minimization using the DFP method until convergence is reached.

The convergence limit used is that both the search direction vector and the step length together must be less than  $10^{-5}$  units.

## RESULTS AND DISCUSSION

As mentioned earlier, a grid search followed by a check for the presence of a 5 $\rightarrow$ 1 hydrogen bond resulted in about 4000 conformations for  $\alpha$ -turns. The minimization has been carried out with each one of these 4000 conformations as starting points. This resulted in 23 unique minimum energy conformations (MECs), spanning the entire conformational space. These are designed as g1 to g23.

The  $(\phi, \psi)$  plots of each of the MECs, at each of three intervening  $\alpha$ -carbon atoms  $C_2^\alpha$ ,  $C_3^\alpha$  and  $C_4^\alpha$ , are shown in Figure 2 (a) to (c). The details of hydrogen bond length and angle of various MECs are given in Table 1. The relative energy values (in kcal/mol) of different MECs with respect to that of the global minimum, g1, and the  $(\phi, \psi)$  values, at the three intervening  $\alpha$ -carbon atoms of these 23 MECs are also given in Table 1.

It is easy to perceive that for each one of the MECs g1 to g23 there exists another with inverse conformation at  $C_2^\alpha$ ,  $C_3^\alpha$  and  $C_4^\alpha$  and these too will have a 5 $\rightarrow$ 1 hydrogen bond. The bond lengths, bond angles, hydrogen and parameters and total energy of a given MEC and its corresponding inverse are the same.

### Classification into Families

A careful look at the  $(\phi, \psi)$  values in Table 1 reveals that not all the minima can be considered as entirely distinct from the conformational viewpoint, although they were arrived at using different starting

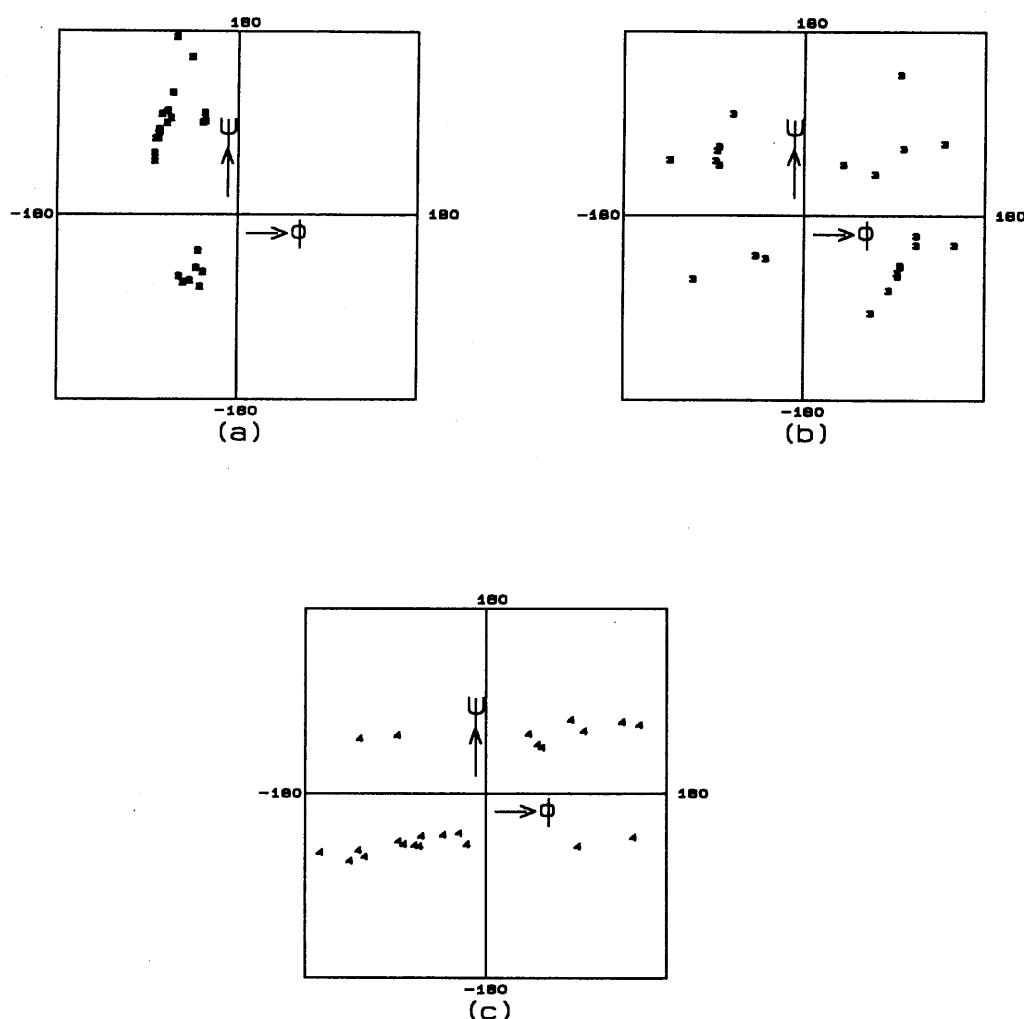


Figure 2  $(\phi, \psi)$  plot of 23 minimum energy conformations of  $\alpha$ -turns: (a)  $(\phi, \psi)$  at  $C_2^\alpha$  (b)  $(\phi, \psi)$  at  $C_3^\alpha$  (c)  $(\phi, \psi)$  at  $C_4^\alpha$ .

points for the minimization. In other words, some of the minima can be considered as local minima occurring in close proximity in the six-dimensional  $(\phi, \psi)$  space, separated by barriers that are not low enough to be surmounted by minimization.

In order to identify and group such close MECs, each MEC is considered as a point in a six-dimensional space (three sets of  $(\phi, \psi)$  at  $C_2^\alpha$ ,  $C_3^\alpha$  and  $C_4^\alpha$ , constitute the coordinates for the six-dimensional space). For any two minima to be conformationally close, it is necessary that the separation in the six-dimensional space (designated as distance) given by:

$$d_{ij} = \sqrt{[(\phi_2^i - \phi_2^j)^2 + (\psi_2^i - \psi_2^j)^2 + (\phi_3^i - \phi_3^j)^2 + (\psi_3^i - \psi_3^j)^2 + (\phi_4^i - \phi_4^j)^2 + (\psi_4^i - \psi_4^j)^2]}$$

should be small. The distances between all possible pairs of MECs are computed. (In view of the cyclic nature of the angular parameters, care is taken to see that the difference between two parameters represents the shortest distance between them, i.e.  $|\Delta\phi|$  or  $|\Delta\psi| \leq 180^\circ$  whenever the difference  $|\Delta\phi|$  or  $|\Delta\psi|$  exceeds  $180^\circ$ , its complement, namely  $360 - |\Delta\phi|$  (or  $|\Delta\psi|$ ) is taken so that it represents the true distance.) In the present study, MECs with distances less than  $100^\circ$  are considered to be conformationally proximal and they are grouped into a class. When this was done, it was found that these 23 MECs fall into 13 classes designated as A, B...M. These are listed in Table 2.

The class A MECs (g1, g3, g7, g13 and g17) have a very similar conformation at  $C_3^\alpha$  and  $C_4^\alpha$  (as seen by the values of  $(\phi_3, \psi_3)$  and  $(\phi_4, \psi_4)$ ). However  $(\phi_2, \psi_2)$  values, though falling in the same quadrant of the

Table 1 Minimum Energy Conformations (MECs) of  $\alpha$ -turns (With Glycyl Backbone) in the Increasing Order of Their Relative Energies. Hydrogen Bond Parameters (Length and Angle) and  $(\phi, \psi)$  Values at the Three Intervening  $\alpha$ -Carbons for These MECs are Also Given

MEC	Hydrogen bond		Relative Energy (kcal/mol)	$\phi, \psi$ values at					
	Length (Å)	Angle (°)							
			$\alpha_2$	$\alpha_3$	$\alpha_4$				
g1	2.97	15.3	0.00	-82.1	81.3	93.4	-51.4	-30.3	-42.4
g2	2.94	10.7	0.25	-56.9	-67.9	-90.6	50.7	-124.1	-64.8
g3	2.93	5.0	0.30	-70.2	92.2	90.4	-58.3	-22.1	-52.9
g4	3.02	26.1	0.54	-35.3	88.9	64.0	-97.2	-91.3	54.9
g5	2.98	15.7	0.56	-38.4	87.9	81.6	-74.9	-129.0	52.4
g6	2.96	19.9	1.03	-40.4	-71.6	-87.9	46.5	81.8	69.6
g7	2.92	13.3	1.07	-48.9	152.3	92.5	-52.1	-74.2	-53.6
g8	2.94	12.7	1.23	-61.0	-61.6	-89.6	61.0	-139.2	-68.5
g9	2.93	6.2	1.41	-44.0	-53.7	-87.7	65.2	39.5	56.0
g10	3.00	24.6	1.55	-35.5	97.0	35.8	47.3	95.2	59.0
g11	2.94	5.0	1.67	-42.2	-36.9	-40.5	-43.6	-90.2	-49.9
g12	2.98	10.9	1.98	-79.4	96.0	67.9	37.5	-129.5	-58.6
g13	2.92	13.8	2.03	-63.5	172.2	90.5	-61.4	-68.6	-55.2
g14	2.96	28.5	2.27	-37.1	-58.4	-73.9	96.6	143.6	-46.4
g15	2.96	18.9	2.43	-67.9	116.8	108.8	-31.4	132.8	67.5
g16	2.99	24.7	2.43	-80.6	79.0	147.1	-30.9	150.4	65.4
g17	2.94	6.9	2.63	-73.1	99.3	108.8	-21.8	-85.0	-53.0
g18	2.95	7.8	2.87	-85.4	71.6	95.7	62.9	-168.1	-60.5
g19	2.98	13.5	3.11	-82.4	72.3	136.5	67.8	48.7	-46.2
g20	2.98	29.5	3.14	-73.8	87.4	93.0	135.1	88.8	-55.2
g21	2.98	29.7	3.35	-50.3	-66.4	-51.2	-40.5	-46.4	-43.9
g22	2.97	16.0	4.21	-85.6	57.6	-112.8	-64.0	-66.8	-45.0
g23	2.98	25.0	4.65	-86.1	50.5	-136.7	52.0	53.1	42.7

$(\phi, \psi)$  map, show some noticeable differences in the value of  $\psi_2$ . The MECs g1, g3 and g17 have their  $\psi_2$  value around  $90^\circ$  while the remaining two, g7 and g13, have their  $\psi_2$  value around  $140^\circ$ . Hence g1, g3 and g17 are grouped into one class, A1, and g7 and g13 into another class, A2. Thus A1 and A2 become sub-classes for the parent class A.

A visual idea of the conformational similarity of MECs of a given class can be had through the  $(\psi, \phi)$  line diagrams. Figure 3 (a to m) shows the line diagrams for the 13 classes of MECs. From the line diagrams the conformational similarity of MECs within a class is apparent. Stereo ORTEP diagrams of lowest MEC of each class are also shown in Figure 3.

As a more convenient representation, each class can be designated by the characteristic conformation at the three intervening  $\alpha$ -carbon atoms, using standard (characterized) regions of the Ramachandran map such as right-handed  $\alpha$ -helical ( $\alpha_R$ ), left-handed  $\alpha$ -helical ( $\alpha_L$ ) or extended (E). The designation of each class is listed in Table 2. Here U denotes

a conformation that does not belong to any of the three standard conformations mentioned above. When we consider MECs with parameters inverse to g1 to g23, it is possible to have 13 more classes (A' to M' inverse to the classes A to M). The family designation will, however, be different (for example  $\alpha_L$  in the place of  $\alpha_R$  and vice versa).

In one of the classes, F, all three sets of  $(\phi, \psi)$  values fall in the right-handed  $\alpha$  helical region. Such a repetition of  $(\phi, \psi)$  values results in an  $\alpha$ -helix. This class is designated as  $\alpha_R\text{-}\alpha_R\text{-}\alpha_R$ . This can be considered to be the basic unit of a  $\alpha$ -helix or ' $\alpha$ -helical  $\alpha$ -turn'.

### Hydrogen Bonding in MECs

The multiple MECs in a given class have been analyzed further, to look for the presence of additional hydrogen bonds, over and above the mandatory 5 $\rightarrow$ 1 hydrogen bond. Within a tetrapeptide segment the hydrogen bonds can be of  $\beta$ -turn type (4 $\rightarrow$ 1 type hydrogen bond) and/or  $\gamma$ -turn type (with

Table 2 MECs of  $\alpha$ -Turns Classified into Different Classes A to M Based on Their Conformational Similarity Along With the Additional Hydrogen Bonds (Other than 5 $\rightarrow$ 1 Type) and the Secondary Structure Identified at Each  $\alpha$ -Carbon Atom

Class	MECs	Additional H-bonds	Secondary structures
A1	g1	$\gamma_2, \gamma_3$	E-U- $\alpha_R$
	g3	$\gamma_3$	E-U- $\alpha_R$
	g17	$\beta_{23}$	E-U- $\alpha_R$
A2	g7	$\beta_{23}, \gamma_3$	E-U- $\alpha_R$
	g13	$\gamma_3$	E-U- $\alpha_R$
B	g2	$\beta_{23}, \gamma_3$	$\alpha_R$ -E-U
	g8	$\gamma_3$	$\alpha_R$ -E-U
C	g4	$\gamma_4$	U-U-U
	g5	$\gamma_3$	U-U-U
D	g6	$\beta_{23}, \gamma_3$	$\alpha_R$ -E- $\alpha_L$
	g9	$\gamma_3$	$\alpha_R$ -E- $\alpha_L$
E	g10	$\beta_{23}$	U- $\alpha_L$ -U
F	g11	$\beta_{23}$	$\alpha_R$ - $\alpha_R$ - $\alpha_R$
	g21	$\beta_{34}$	$\alpha_R$ - $\alpha_R$ - $\alpha_R$
G	g12	$\beta_{23}$	E- $\alpha_L$ -U
	g18	$\gamma_2$	E- $\alpha_L$ -U
H	g14	$\beta_{34}$	$\alpha_R$ -E-U
I	g15	$\beta_{23}$	E-U-U
	g16	$\beta_{23}, \gamma_2$	E-U-U
J	g19	$\gamma_2$	E-U- $\alpha_L$
K	g20	$\gamma_4$	E-U-U
L	g22	$\gamma_2$	U- $\alpha_R$ - $\alpha_R$
M	g23	$\gamma_2$	U-U- $\alpha_L$

$\alpha_R$ , right-handed  $\alpha$ -helical conformation;  $\alpha_L$ , left-handed E, extended conformation;  $\gamma_i \Rightarrow 3 \rightarrow 1$  type hydrogen bond at  $C_i^{\beta}$   
 $\beta_{mn} \Rightarrow 4 \rightarrow 1$  type hydrogen bond across  $C_m^{\alpha}$  and  $C_n^{\alpha}$ .

3 $\rightarrow$ 1 type hydrogen bond). The 4 $\rightarrow$ 1 type hydrogen bond can be any one of the following types: (i) N<sub>4</sub>-H<sub>4</sub>...O<sub>1</sub> ( $\beta_{23}$ ) or (ii) N<sub>5</sub>-H<sub>5</sub>...O<sub>2</sub> ( $\beta_{34}$ ). Similarly, the 3 $\rightarrow$ 1 type hydrogen bond can be any one of the following types: (i) N<sub>3</sub>-H<sub>3</sub>...O<sub>1</sub> ( $\gamma_2$ ); (ii) N<sub>5</sub>-H<sub>5</sub>...O<sub>3</sub> ( $\gamma_4$ ), (iii) N<sub>4</sub>-H<sub>4</sub>...O<sub>2</sub> ( $\gamma_3$ ); (iv) N<sub>5</sub>-H<sub>5</sub>...O<sub>3</sub> ( $\gamma_4$ ).

A schematic representation of the above-mentioned possibilities of these hydrogen bonds is shown as dotted lines in Figure 4. It can be seen that the arrangements shown in Figure 4(b) and 4(c) leads to bifurcated hydrogen-bonding system. When the MECs are examined for the above types of hydrogen bonds they did show such patterns. The details are given in Table 2. From the table it can be noticed that (i) all the MECs have one more hydrogen bond in addition to the 5 $\rightarrow$ 1 hydrogen bond and (ii) the different MECs in a given class have different hydrogen-bonding patterns.

### Feasibility of Accommodating Non-Gly Residues in MECs

The MECs of  $\alpha$ -turns discussed in the previous section have been obtained with glycol residues in the backbone. As mentioned earlier, Gly being the smallest amino acid offers the least stereochemical restrictions and the MECs obtained are dictated by backbone atomic interactions and hydrogen bonding effects alone. Hence, these MECs can be regarded as basic prototypes for  $\alpha$ -turns. Substitution by other non-Gly residues at any of the three positions will result in the elimination of some of the MECs on stereochemical grounds. In naturally occurring molecules, such as peptides and proteins, any of the 20 amino acids can participate in such  $\alpha$ -turns [1]. Therefore it is worth-while to examine these MECs for the feasibility of accommodating non-Gly residues. The result of such a study is given in this section.

The 20 amino acids can be divided into three categories: (i) Gly, which lacks a C <sup>$\beta$</sup>  and which has maximum conformational flexibility; (ii) Pro, which is an imino acid, with a pyrrolidine ring and has least conformational flexibility. In addition, it is known that Pro has a higher potential to occur at the central positions of reverse turns ( $\beta$ -turns) [10–13]; (iii) all the other non-Gly, non-Pro amino acids. Alanine has been taken as the representative for all these. The MECs and their inverses are examined to find out the possibility of accommodating Ala/Pro.

*Accommodation of L-Alanyl Residue.* The feasibility of accommodating Ala at any one or more C <sup>$\alpha$</sup> s in the central locations of the  $\alpha$ -turn MEC can be readily obtained by superimposing line diagrams of the MECs on the L-Ala Ramachandran map [15]. As an example, the superposition of MECs of the class A2 on the Ramachandran map for L-Ala is shown in Figure 5. From the figure it can be easily deduced

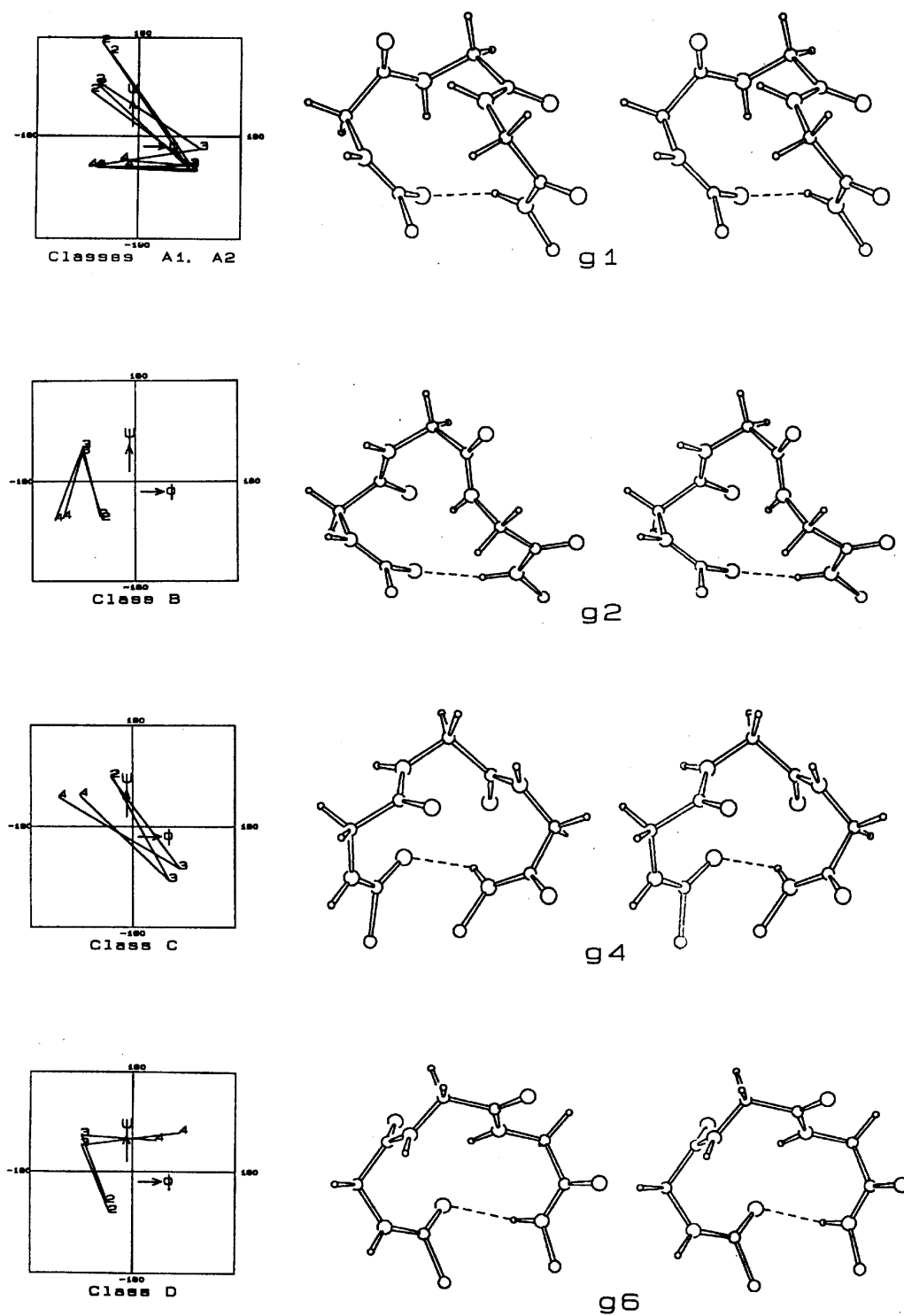


Figure 3 Line diagrams of member MECs of different classes. Stereo ORTEP pairs of lowest MEC of each class are shown on the right.

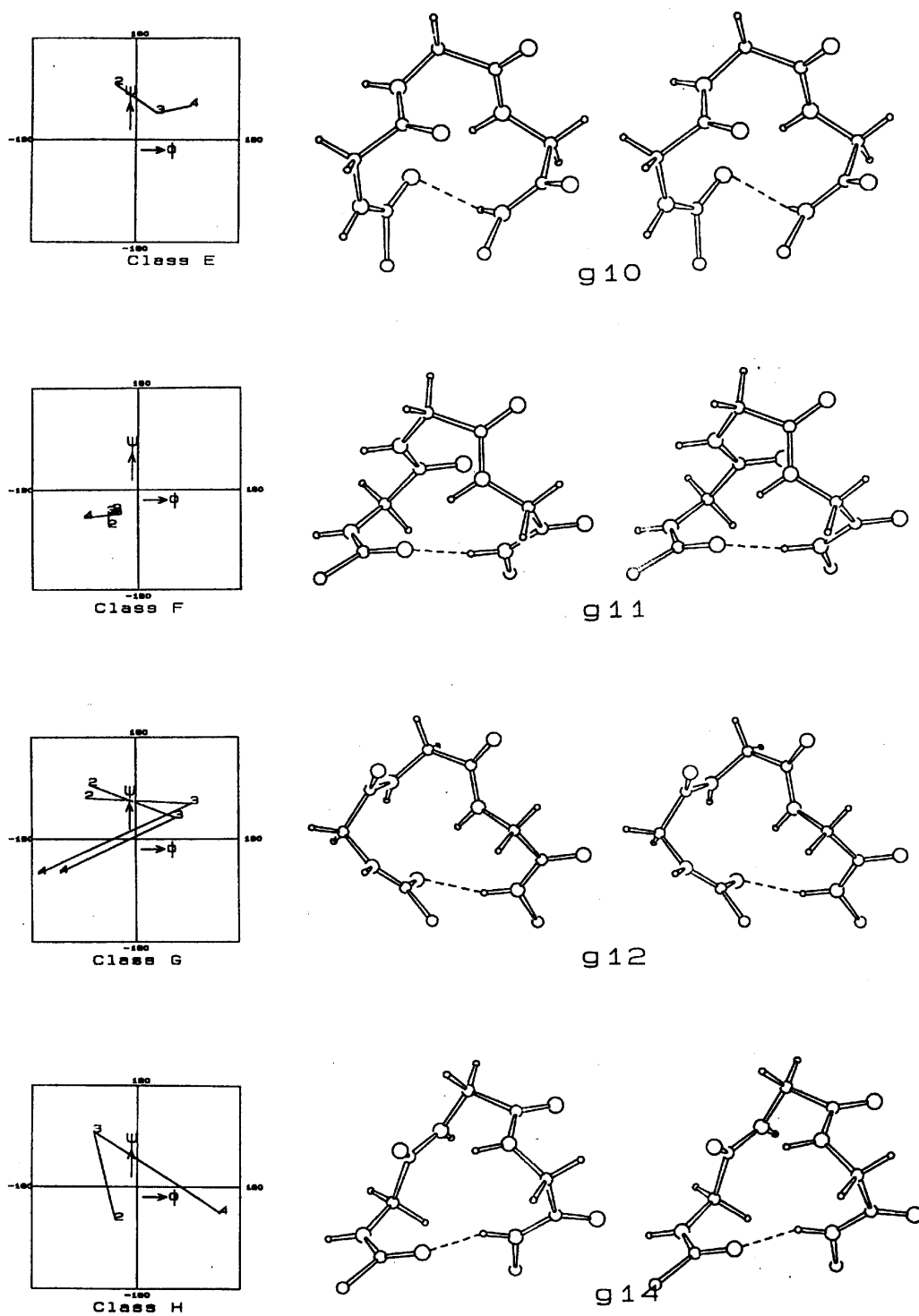


Figure 3. (continued)

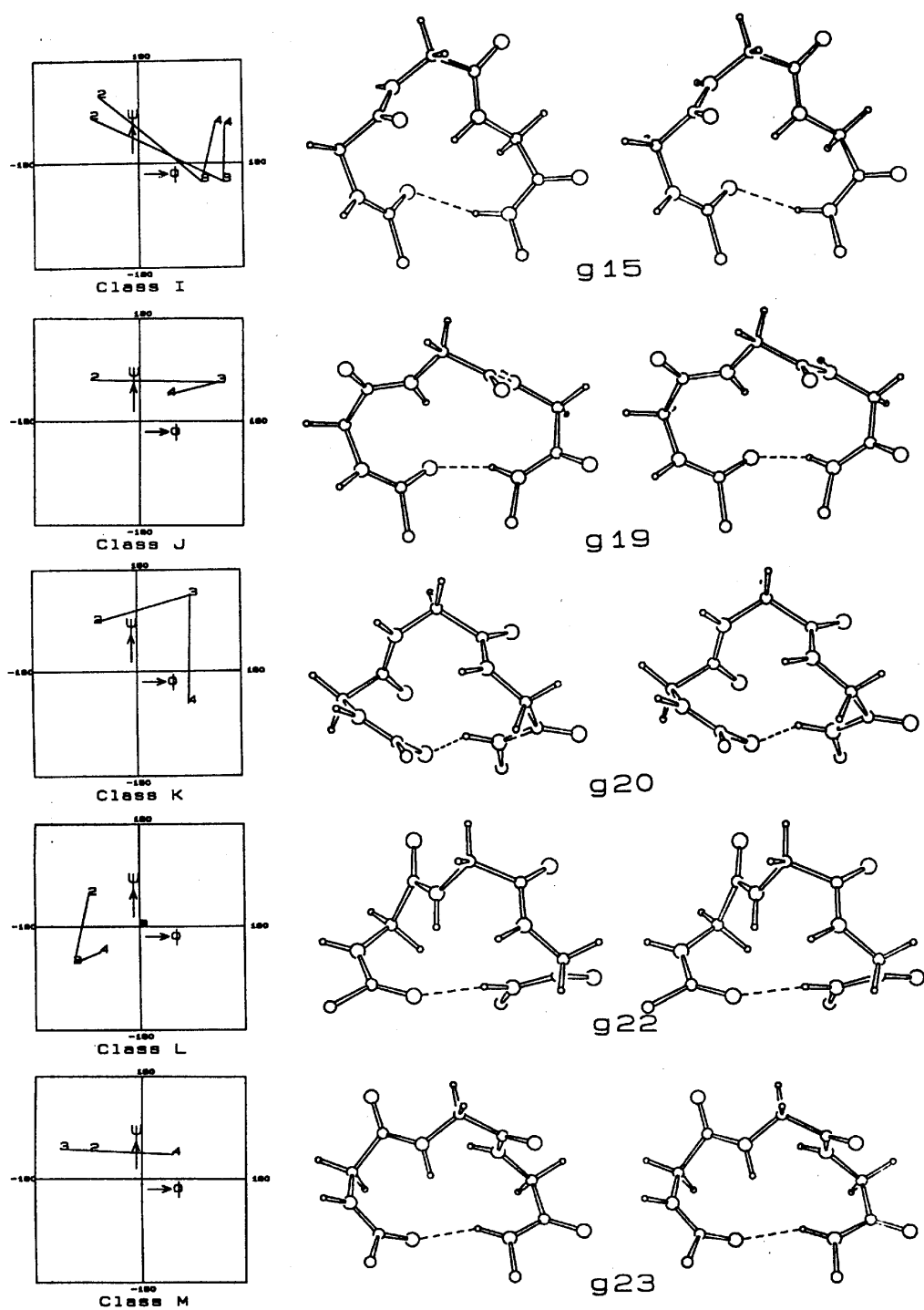


Figure 3. (continued)



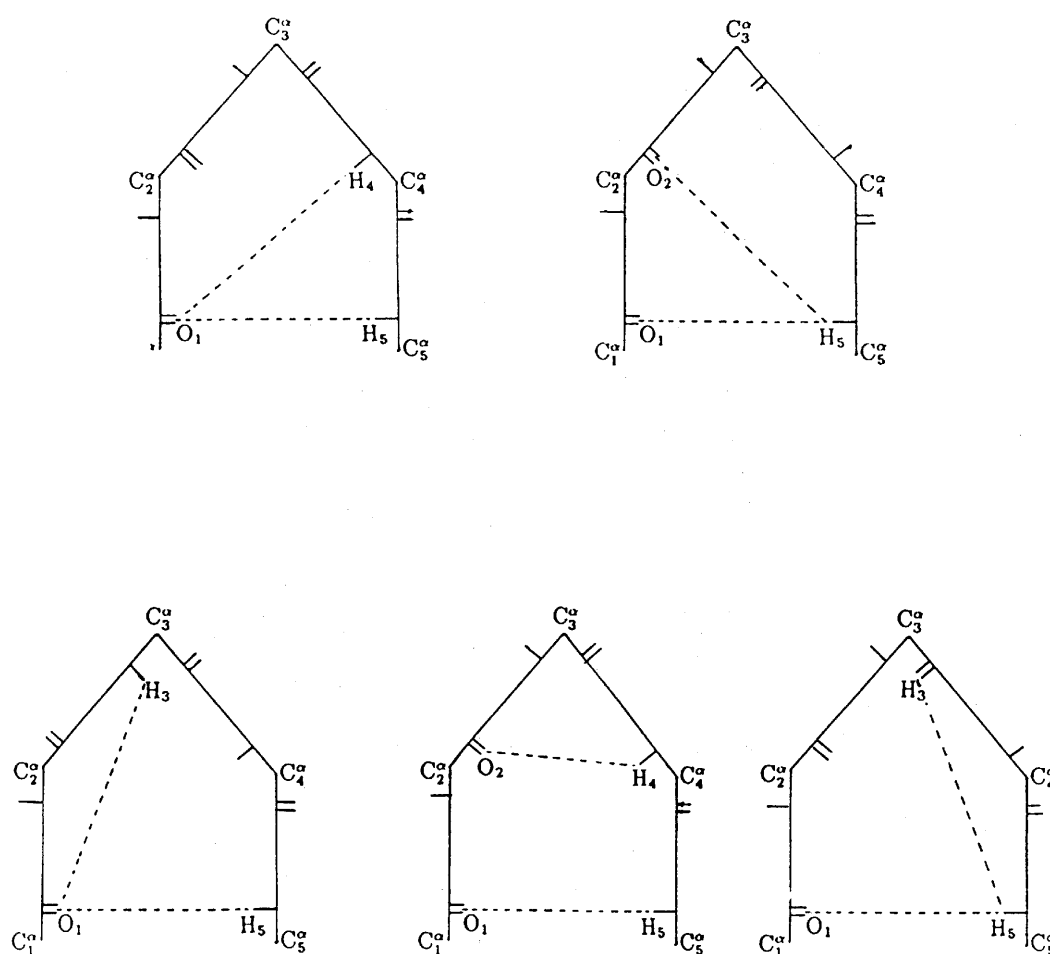


Figure 4 Schematic diagrams of different types of hydrogen bonds (in addition to the 5 $\rightarrow$ 1 hydrogen bond) possible in the system of a four linked peptide units.

that while the conformation at  $C_2^\alpha$  and  $C_4^\alpha$  can accommodate L-Ala, that at  $C_3^\alpha$  can accommodate only Gly residue.

*Accommodation of Pro.* To accommodate Pro at any  $C^\alpha$  site, the conformation should be suitable for pyrrolidine ring closure in addition to its suitability for accommodating Ala. A value for  $\phi$  of around  $-60^\circ$  is ideal for L-proline ring closure. But prolyl residues in peptide and proteins show a deviation in these  $\phi$  values [14–16]. In fact pyrrolidine ring closure can be achieved with minimal distortion when the value of  $\phi$  is  $-60^\circ$ . Hence, in any MEC, the  $\alpha$ -carbon atom is considered to be suitable for the accommodation of Pro, if it is suitable for Ala and if  $-90^\circ \leq \phi \leq -40^\circ$ .

Tables 3 and 4 give the accommodating capability of different MECs for L-Ala and/or L-Pro. The absence of 'A' or 'P', denoted by (-), means only glycyl residue can be accommodated. The number of possible sequences (of Gly, Ala and Pro) for a given

prototype MEC is given in Tables 3 and 4. It is possible to have 371 (249 + 122) sequences that are geometrically and sterically possible. From the tables it can be seen that at least three sequences are possible for a given prototype MEC conformation (parameter  $n$ ). At the other extreme, the MECs g11 and g21 (belonging to the class F) can accommodate A, G and P at all the three positions leading to a maximum of 27 possibilities.

In order to ascertain whether these minima, after incorporating the corresponding combination of Gly/non-Gly residues are energetically favourable, systematic energy minimization studies were also carried out. The convergence was reached in 321 out of 371 combinations. However in ten cases the 5 $\rightarrow$ 1 hydrogen bond was found broken after minimization. (The complete details of the sequence combination and  $(\phi, \psi)$  parameters of the MECs are not reported here and can be had from the authors on request.) Except in 24 sequence combinations,

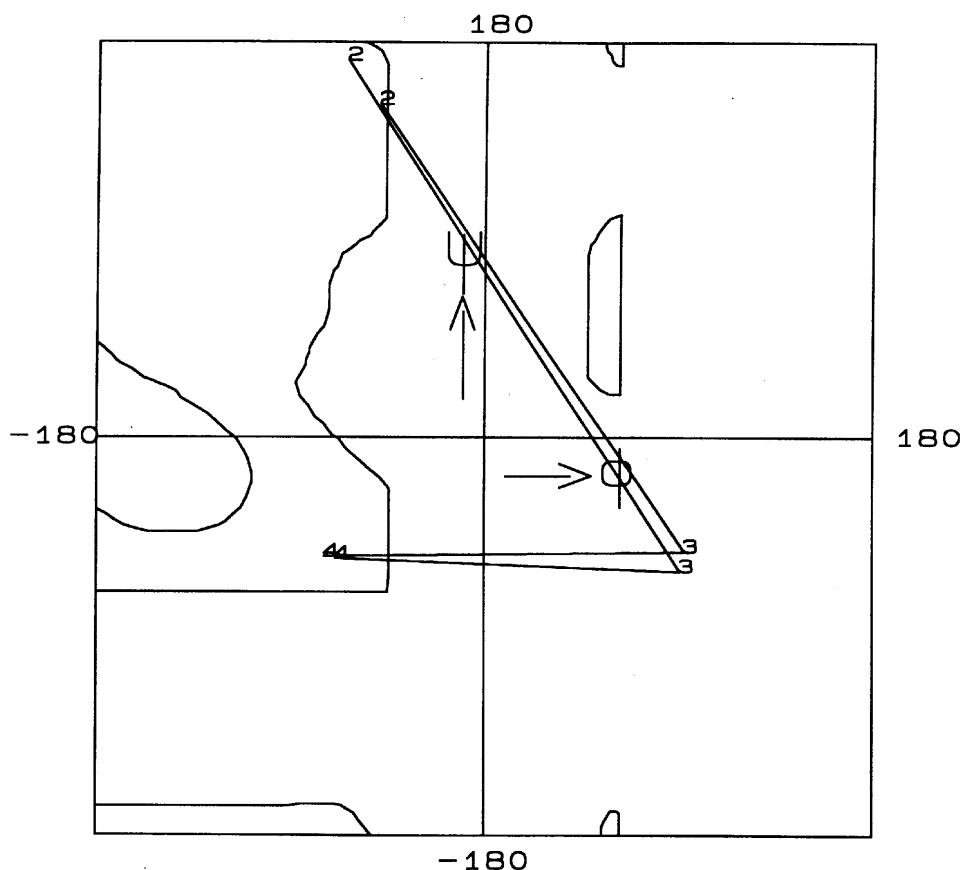


Figure 5 Line diagrams of member MECs of class A2 superimposed on Ramachandran map for L-Ala.

minimization always led to an MEC with the closest prototype being the same as the starting prototype. These 24 combinations are also given in Tables 3 and 4.

#### Comparison of MECs with Examples Observed in Proteins

In order to compare the present results with the  $\alpha$ -turns observed in proteins, the Brookhaven Protein Data Bank (PDB) [17] has been used. For this purpose, a set of 250 proteins with resolution 2 Å or better are chosen such that sequence homology between any two members of the set is not greater than 40%. Also in the case of proteins having multiple chains with the same sequence, only one chain has been considered. Protein codes of these 250 proteins are given in [19]. The  $\alpha$ -turn in a protein is identified when 5→1 hydrogen bond occurs in a four linked peptide segment and does not form part of an  $\alpha$ -helix. The criteria used for this purpose is a relaxed one compared with that mentioned in an earlier section, (a) the hydrogen bond length N...O

must be less than 3.6 Å and (b) the hydrogen bond angle H-N...O  $\leq 40^\circ$ . Amide hydrogens were fixed geometrically in order to calculate the H-N...O angle.

There are 216 examples of  $\alpha$ -turns that could be identified occurring in 122 proteins [17]. In fact this is an updated set of the one reported in an earlier study from this group [1] which had 96 examples.

The observed examples were grouped into 14 families (with the largest number of examples occurring as  $\alpha$ -helical  $\alpha$ -turns). The average ( $\psi$ ,  $\phi$ ) values for each family can be compared with the MECs obtained in this study. It should, however, be kept in mind that at times it may not be possible to get good agreement between these MECs and the observed examples of  $\alpha$ -turns in proteins, since there are other interactions which can easily modify these conformations. In additions, bulkier side groups can also play a role in bringing about deviations from the theoretical prototypes.

1. For any observed example, the sequence of the residues at the three middle  $\alpha$ -carbon atoms are looked into and the relevant theoretical MECs with

Table 3 Accommodating Capability at the Three C $^{\alpha}$  Positions in Different MECs (Direct) and the Number of Combinatorial Possibilities ( $n$ ) (A, L-Ala; P, L-Pro; (-) Indicates That Only Glycyl Residue Can Be Accommodated. Whenever Energy Minimization Results in a Minima Different From the Starting Prototype, the Sequence and the Prototype to Which it Converges are Given in the Last Column

Class	Direct	At			$n$	Sequences converging to different prototype minima
	MEC	C $_2^{\alpha}$	C $_3^{\alpha}$	C $_4^{\alpha}$		
A1	g1	A P	-	A	6	
	g3	A P	-	A	6	
	g17	A P	-	A	6	3 AGA(g12) AGG(g7) GGA(g12)
A2	g7	A P	-	A P	9	2 AGP(g1) GGP(g1)
	g13	A P	-	A P	9	3 AGA(g3) AGG(g3) GGP(g1)
B	g2	A P	A P	A	18	
	g8	A P	A P	A	18	2 AAA(g2) APG(g14)
C	g4	A	-	A	4	
	g5	A	-	A	4	
D	g6	A P	A P	A	18	4 AAA(g9) AGA(g9) GPA(g9) PAA(g9)
	g9	A P	A P	A	18	
E	g10	A	A	-	4	
F	g11	A P	A P	A P	27	
	g21	A P	A P	A P	27	
G	g12	A P	A	A	12	
	g18	A P	A	A	12	8 AAA(g12) AAG(g12) AGA(g12) GAA(g12) GGA(g12) AGG(g12) GAG(g12) PAA(g12)
H	g14	A	A P	-	6	
I	g15	A P	-	-	3	
	g16	A P	-	-	3	
J	g19	A P	-	A	6	
K	g20	A P	-	-	3	
L	g22	A P	A	A P	18	
M	g23	A P	A	A	12	
<b>Total</b>					<b>249</b>	

Table 4 Accommodating Capability at the Three C<sup>α</sup> Positions in Different MECs (Inverse) and the number of Combinatorial Possibilities

Class	Inverse		At			<i>n</i>	Sequences converging to different prototype minima
	MEC	C <sub>2</sub> <sup>α</sup>	C <sub>3</sub> <sup>α</sup>	C <sub>4</sub> <sup>α</sup>			
A1'	i1	-	A P	A	6		
	i3	-	A P	A	6		
	i17	-	A	A	4		
A2'	i7	A	A P	A	12		
	i13	A	A P	A	12		
B'	i2	A	-	-	2		
	i8	A	-	-	2		
C'	i4	-	A P	-	3		
	i5	-	A P	-	3	1 GPG(g14)	
D'	i6	A	-	A P	6	1 GGP(i9)	
	i9	A	-	A P	6		
E'	i10	-	A	A P	6		
F'	i11	A	A	A	8		
	i21	A	A	A	8		
G'	i12	-	A P	-	3		
	i18	-	A P	-	3		
H'	i14	A	-	A	4		
I'	i15	-	A	A	4		
	i16	-	A	A	4		
J'	i19	A	A	A P	12		
K'	i20	-	-	A P	3		
L'	i22	-	-	A	2		
M'	i23	-	-	A P	3		
<b>Total</b>					<b>122</b>		

which the comparison should be made are identified. For example, if the sequence is Gly-Asp-Pro, the relevant theoretical MECs are those that have been worked out for the sequence Gly-Ala-Pro (GAP) (in other words, all the non-Gly and non-Pro residues are approximated to Ala residue).

2. The similarity of an observed example with any MEC is quantified by the six parameter distance

calculation (this is similar to the method used to classify the observed examples as well as theoretical MECs into families). The upper limit cut-off of 100° for the 'distance', in the six parameter space, has been used to pick out MECs that are conformationally close to the observed conformation. The class of that MEC is considered to be one that is closest to the family to which the observed conformation

Table 5 Comparison of the Theoretical MEC Classes with Observed Families in Proteins<sup>a</sup>. The Average values at the Three Intervening  $\alpha$ -Carbon Atoms of the Various Observed Families are also Included

Family of the observed examples	Average ( $\phi$ , $\psi$ ) values at						No. of observed examples	No. for which close MEC could be found	Theoretical MEC families that agree with observation <sup>b</sup>
	$\alpha_2$	$\alpha_3$	$\alpha_4$						
F1	-63	-25	-77	-33	-103	-20	144	142	B(4), F(136), L(2)
F2	-59	131	82	-7	-114	-51	21	21	A1(6), A2(8), G(7)
F3	55	25	65	11	-119	-35	9	5	D'
F4	-60	149	64	32	60	39	7	2	E
F5	-52	-53	-76	111	123	-39	7	4	B(2), H(2)
F6	57	-137	-97	74	64	36	6	5	A1'
F7	57	49	64	51	67	31	6	6	F'
F8	58	-150	-86	-12	-102	-29	6	4	E'(1), I'(3)
F9	-76	77	171	-54	-83	-33	3	1	A2
F10	-60	-24	-91	6	96	37	2	2	D
F11	51	55	68	-46	-141	5	2	0	
F12	-145	102	-63	-48	-53	-49	1	0	
F13	-58	126	43	42	-129	-46	1	1	G
F14	-30	-64	-162	61	-126	-54	1	1	B
							<b>216</b>	<b>194</b> (~90%)	

<sup>a</sup>See text for the method of identifying close MEC.

<sup>b</sup>Whenever there are multiple families, the figures in the parentheses show the distribution of the examples among the different families.

belong. The consolidated results are given in Table 5. It is very satisfying to note that overall 90% of the observed examples (194 out of 216) can be associated with one or other of the theoretical MECs.

## CONCLUSIONS

The present study shows that, like  $\beta$  and  $\gamma$  turns, it is possible to identify a novel turn feature, *viz.* the  $\alpha$ -turn with characteristic ( $\phi$ ,  $\psi$ ) values at the three intervening  $\alpha$ -carbons in a system of all four *trans* tetrapeptide segment, with a 5 $\rightarrow$ 1 type hydrogen bond between the oxygen of first peptide unit and the amide hydrogen of fourth peptide unit. It is possible to have 15 (plus the inverse of these) prototype classes of  $\alpha$ -turns with glycyl backbone, in the whole conformational space. The existence of a greater number of classes for an  $\alpha$ -turn as compared to  $\beta$ -turn is due to the larger degree of freedom for  $\alpha$ -turn. (While the definition of a  $\beta$ -turn requires only two sets of ( $\psi$ ,  $\phi$ ) values the  $\alpha$ -turn requires three sets.) The fact that at least some examples of different  $\alpha$ -turns have been observed in proteins shows that these  $\alpha$ -turns can be a characteristic feature of protein conformation, through playing a minor role in protein folding.

The smaller number of occurrences of isolated  $\alpha$ -turns can be qualitatively attributed to the basic nature of successive occurrence of these turns, resulting in more stable  $\alpha$ -helices, abounding in proteins.

## REFERENCES

1. D. V. Nataraj, N. Srinivasan, R. Sawdhamini and C. Ramakrishnan (1995). *Curr. Sci. (India)* 69, 434–447.
2. C. M. Venkatachalam (1968). Stereochemical criteria for polypeptides and proteins. Conformation of a system of three linked peptide units. *Biopolymers* 6, 1425–1436.
3. R. Chandrasekaran, A. V. Lakshminarayanan, U. V. Pandya and G. N. Ramachandran (1974). Conformation of the LL and LD hairpin bends with internal hydrogen bonds in peptides and proteins. *Biochim. Biophys. Acta* 303, 14–27.
4. R. B. Corey and L. Pauling (1953). Dimensions of polypeptide chains. *Proc. R. Soc. (London)* B141, 10–20.
5. G. N. Ramachandran and V. Sasisekharan (1968). Conformation of polypeptides and proteins. *Adv. Protein Chem.* 23, 283–437.
6. P. K. C. Paul and C. Ramakrishnan (1985). Studies on hydrogen bonds Part V – Hydrogen bonding in energy minimization studies of peptide. *J. Biomol. Structure Dynamics* 2, 879–989.

7. W. Murray in: *Numerical Methods for Unconstrained Optimization*, Academic Press, London 1972.
8. S. R. Niketic and K. Rasmussen in: *Lecture Notes in Chemistry*, Vol. 3, Springer Verlag 1977.
9. R. Fletcher in: *Practical Methods of Optimization*, Wiley, New York 1980.
10. P. Y. Chou and G. D. Fasman (1977).  $\beta$ -Turns in Proteins. *J. Mol. Biol.* 115, 135–175.
11. J. S. Richardson (1981). The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* 34, 167–339.
12. G. D. Rose, L. M. Gierasch and J. A. Smith (1985). Turns in peptides and proteins. *Adv. Protein Chem.* 37, 1–109.
13. C. M. Wilmot and J. M. Thornton (1988). Analysis and prediction of the different types of  $\beta$ -turn in proteins. *J. Mol. Biol.* 203, 221–232.
14. V. N. Balaji, PhD Thesis Molecular Biophysics Unit, Indian Institute of Science, Bangalore 1980.
15. M. J. O. Anteunis and J. J. M. Sleekx in: *Molecular Structure and Energetics*, p. 189, Vol. 4, J. F. Leibman and A. Greenberg Eds., VCH Publishers Inc. (1987).
16. J. S. Richardson and D. C. Richardson (1989). The *de novo* design of protein structures. *Trends Biochem. Sci.* 14, 304–309.
17. D. V. Nataraj, PhD Thesis Molecular Biophysics Unit, Indian Institute of Science, Bangalore 1996.
18. F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice Jr, J. R. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi (1977). The Protein Data Bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112, 535–542.
19. K. Gunasekaran, C. Ramakrishnan and P. Balaram (1997).  $\beta$ -Hairpins in proteins revisited: Lessons for *de novo* design. *Prot. Eng.* (in press).