# Protein secondary structure

**Regular, recurring arrangements in space of adjacent amino acid residues in a polypeptide chain.**

**Maintained by hydrogen bonds between amide hydrogen and carbonyl oxygen of the peptide backbone.**

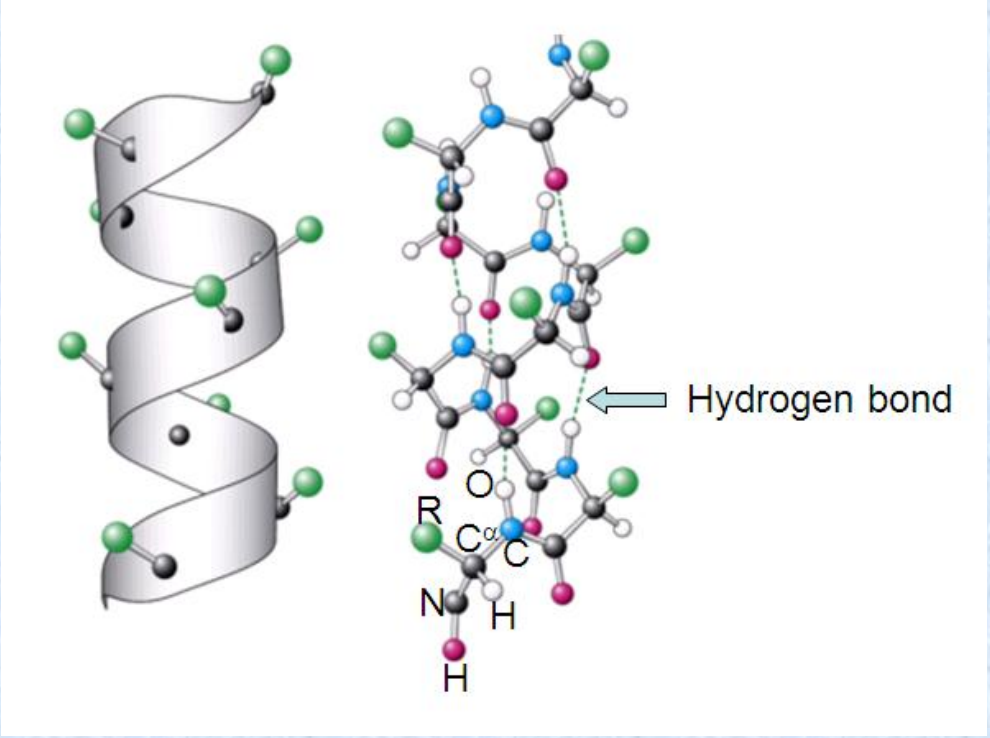
**1. α-helices**

**2. β-strands (sheets)**

**3. Turns (bends)**

**4. Coil (irregular)**

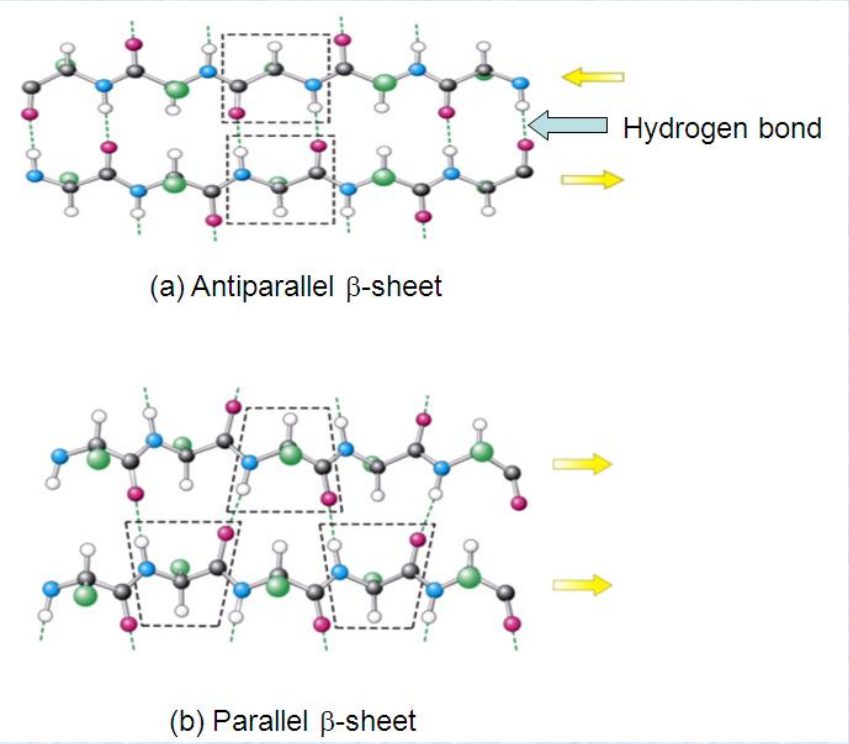
# Secondary structures: α-helix

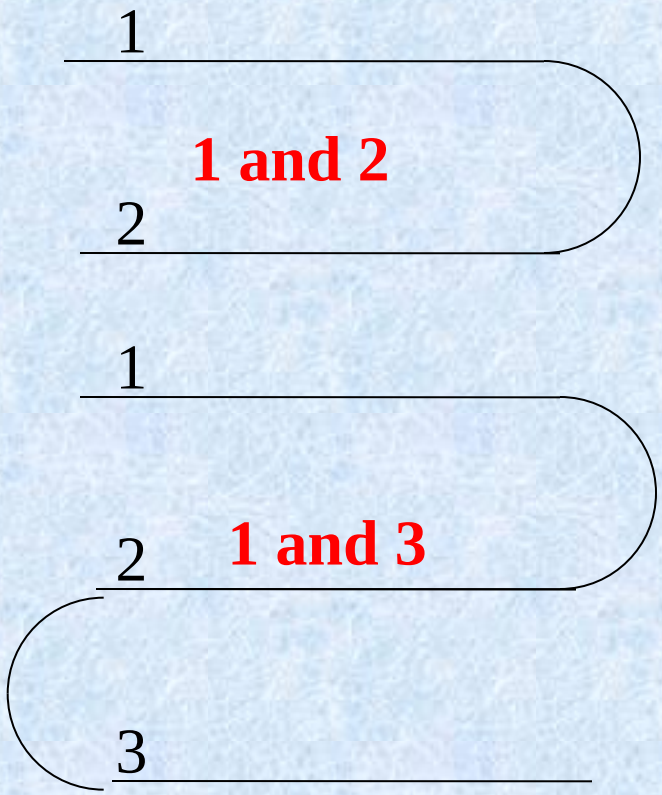



The α-helix is stabilized by hydrogen bonds between the CO and the NH groups of the main chain (i and i+4). 3.6 residues/turn; 5.4 Å/turn (repeating unit of α-helix)

Residues are closely packed.

# Secondary structures: β-strand




**β-strand is fully extended and loosely packed; they can hydrogen bond together to form β-sheets (parallel or anti parallel) depending on the direction.**

# Turns

One-third of all residues in globular proteins are contained in turns that serve to reverse the direction of the polypeptide chain.

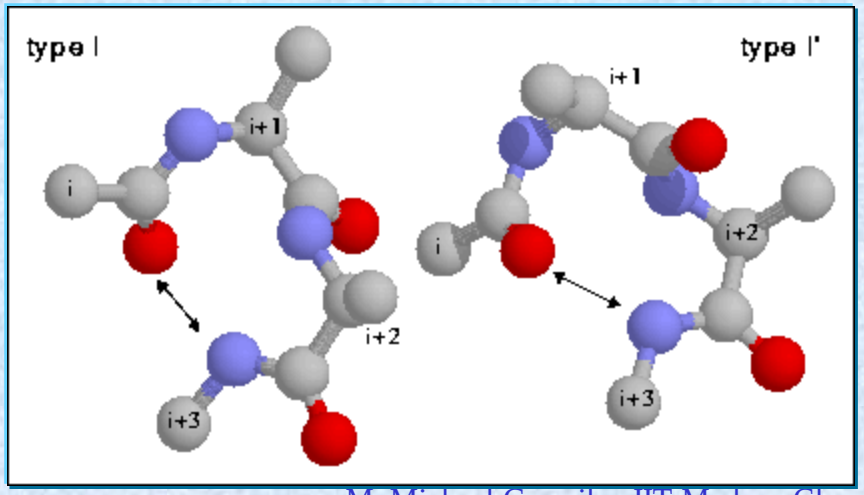
Turns contain a hydrogen bond between the carbonyl oxygen of residue i and the amide nitrogen of i+3.

These three types of turns are designated I, II, and III.

Type I turns occur most frequently (2-3 times more frequently than type II).

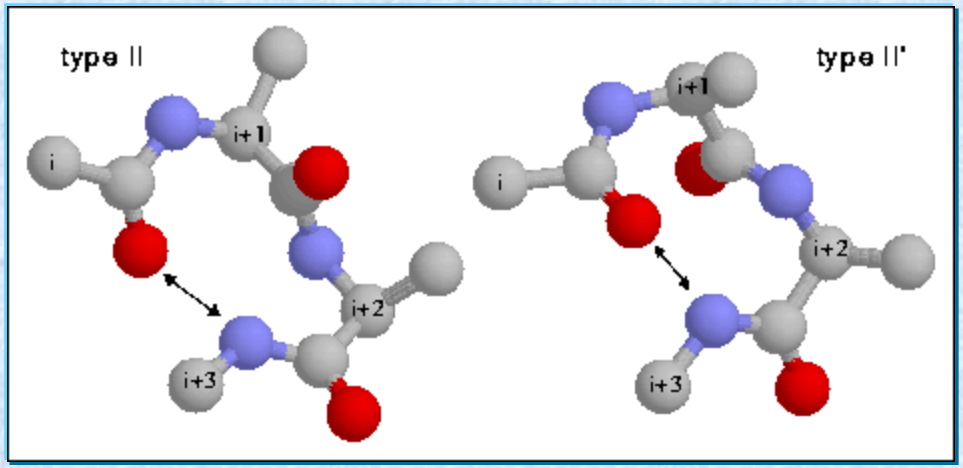
The backbone dihedral angles of residue are (-60, -30) and (-90, 0) of residues i+1 and i+2 of the type I turn.

Proline is often found in position i+1 in type I turns as its phi angle is restricted to -60.



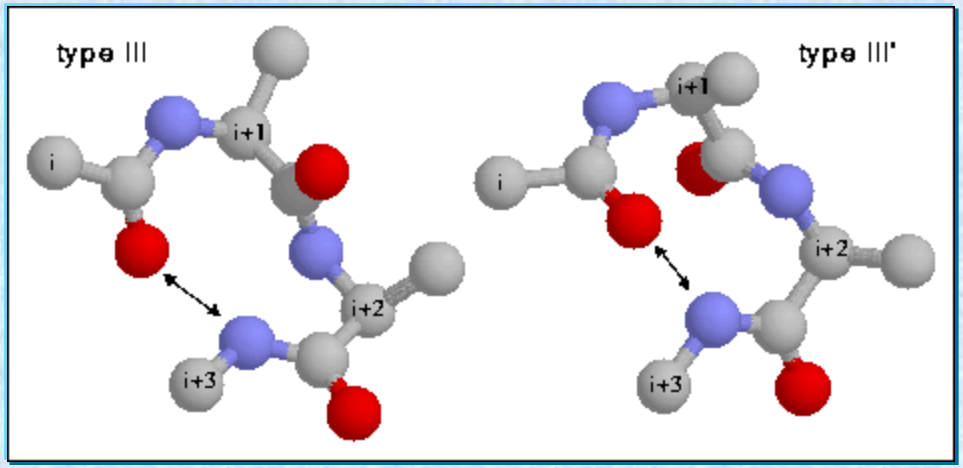
# Turns

**The backbone dihedral angles of residue are (-60, 120) and (80, 0) of residues i+1 and i+2, respectively of the type II turn.**

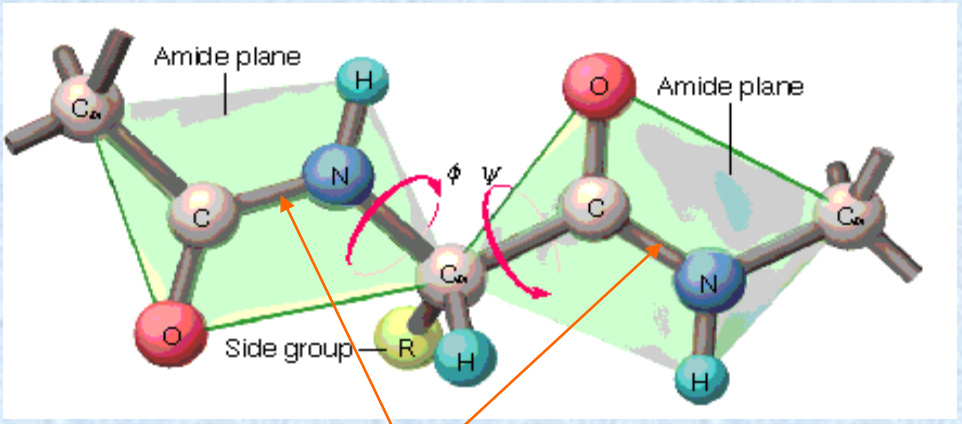



**Glycine is favored in this position in the type II' as it requires a positive phi value.**

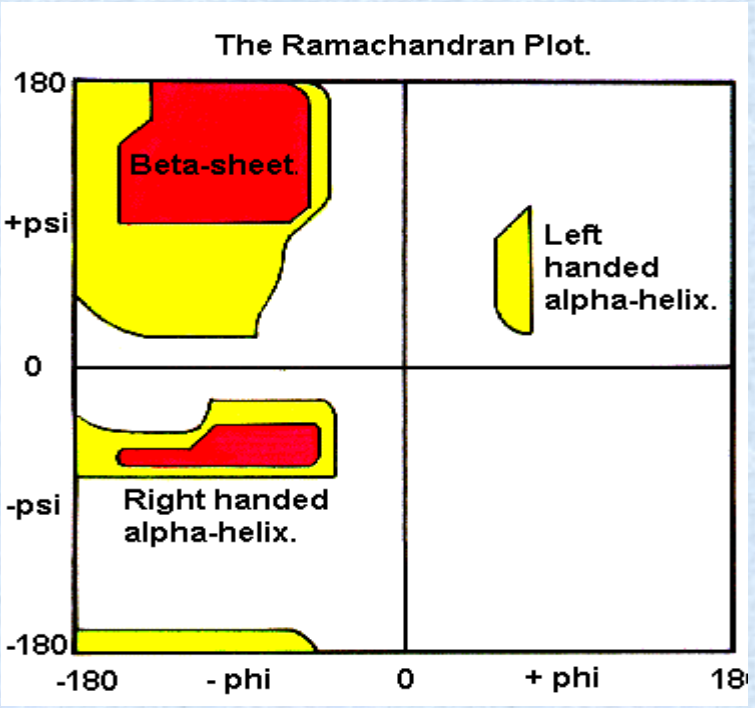
**The backbone dihedral angles of residue are (-60, -30) and (-60, -30) of residues i+1 and i+2, of the classical type III turn.**



# Ramachandran plot




N-$C_\alpha$ and $C_\alpha$-C bonds are free to rotate.
These rotations are represented by the torsional angles, Φ and Ψ, respectively.




G N Ramachandran used computer models of small polypeptides to systematically vary Φ and Ψ for finding stable conformations.
Atoms were treated as hard spheres (dimensions -> van der Waals radii).
Φ and Ψ angles that cause spheres to collide -> sterically disallowed conformation

Red: Allowed; White: Disallowed

# Torsional angles

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ATOM | 27 | CG | GLU | A | 4 | 7.583 | 14.248 | 24.732 | 1.00 | 14.95 | C |
| ATOM | 28 | CD | GLU | A | 4 | 7.577 | 15.734 | 24.340 | 1.00 | 15.89 | C |
| ATOM | 29 | OE1 | GLU | A | 4 | 6.511 | 16.344 | 24.159 | 1.00 | 17.02 | O |
| ATOM | 30 | OE2 | GLU | A | 4 | 8.572 | 16.043 | 23.730 | 1.00 | 15.37 | O |
| ATOM | 31 | N | GLY | A | 5 | 7.687 | 11.127 | 22.185 | 1.00 | 12.43 | N |
| ATOM | 32 | CA | GLY | A | 5 | 8.655 | 10.062 | 21.811 | 1.00 | 14.67 | C |
| ATOM | 33 | C | GLY | A | 5 | 9.176 | 10.126 | 20.356 | 1.00 | 16.35 | C |
| ATOM | 34 | O | GLY | A | 5 | 10.347 | 9.825 | 20.076 | 1.00 | 11.57 | O |
| ATOM | 35 | N | GLU | A | 6 | 8.315 | 10.552 | 19.440 | 1.00 | 13.50 | N |
| ATOM | 36 | CA | GLU | A | 6 | 8.800 | 10.689 | 18.054 | 1.00 | 12.50 | C |
| ATOM | 37 | C | GLU | A | 6 | 9.632 | 11.962 | 17.942 | 1.00 | 13.96 | C |
| ATOM | 38 | O | GLU | A | 6 | 10.617 | 11.986 | 17.195 | 1.00 | 14.23 | O |
| ATOM | 39 | CB | GLU | A | 6 | 7.614 | 10.784 | 17.076 | 1.00 | 13.07 | C |

(x1,y1,z1)
(x2,y2,z2)
(x3,y3,z3)

## Dihedral Angles

### Surface One

| | x | y | z |
|---|---|---|---|
| Point 1 | 9.1760 | 10.126 | 20.3560 |
| Point 2 | 8.3150 | 10.5520 | 19.4400 |
| Point 3 | 8.8000 | 10.6890 | 18.0540 |

Distances

| | |
|---|---|
| 1-2 | 1.32734810807 |
| 2-3 | 1.47478473005 |
| 3-1 | 2.39948932066 |

Equation of Plane Ax+By+Cz+D=0

-0.46494399999999 x + -1.637606 y + -0.324567 z + 27.4556103519997 = 0

### Surface Two

| | x | y | z |
|---|---|---|---|
| Point 1 | 8.3150 | 10.5520 | 19.4400 |
| Point 2 | 8.8000 | 10.6890 | 18.0540 |
| Point 3 | 9.6320 | 11.9620 | 17.9420 |

Distances

| | |
|---|---|
| 1-2 | 1.47478473005 |
| 2-3 | 1.52489245522 |
| 3-1 | 2.44266104893 |

Equation of Plane Ax+By+Cz+D=0

1.749034 x + -1.09883199999999 y + 0.50342100000000 z + -12.734846686000 = 0

**Dihedral Angle =** 1.34556071239086 radians

77.094949898613 degrees

$$A = \begin{vmatrix} 1 & y1 & z1 \\ 1 & y2 & z2 \\ 1 & y3 & z3 \end{vmatrix} \quad B = \begin{vmatrix} x1 & 1 & z1 \\ x2 & 1 & z2 \\ x3 & 1 & z3 \end{vmatrix} \quad C = \begin{vmatrix} x1 & y1 & 1 \\ x2 & y2 & 1 \\ x3 & y3 & 1 \end{vmatrix} \quad D = - \begin{vmatrix} x1 & y1 & z1 \\ x2 & y2 & z2 \\ x3 & y3 & z3 \end{vmatrix}$$

## Dihedral Angle Calculator

A dihedral angle is the angle between two planes.

To calculate this angle, you can follow these steps:

1. **Calculate the equation for each plane**. It will be in the form:

$$Ax + By + Cz + D = 0$$

2. Then, knowing the equation of the two planes, you can **calculate the dihedral angle**:

$$\cos\alpha = \frac{A_1A_2 + B_1B_2 + C_1C_2}{\sqrt{A_1^2 + B_1^2 + C_1^2}\sqrt{A_2^2 + B_2^2 + C_2^2}}$$

# Check whether the residues 60-69 form an α-helix (4MBN)

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ATOM | 483 | N | ASP | A | 60 | 30.554 | 28.872 | -1.439 | 1.00 | 13.41 |
| ATOM | 484 | CA | ASP | A | 60 | 29.251 | 29.577 | -1.346 | 1.00 | 13.47 |
| ATOM | 485 | C | ASP | A | 60 | 28.082 | 28.593 | -1.208 | 1.00 | 10.85 |
| ATOM | 486 | O | ASP | A | 60 | 27.129 | 28.854 | -0.468 | 1.00 | 10.89 |
| ATOM | 487 | CB | ASP | A | 60 | 29.115 | 30.488 | -2.563 | 1.00 | 11.63 |
| ATOM | 488 | CG | ASP | A | 60 | 27.959 | 31.474 | -2.407 | 1.00 | 11.34 |
| ATOM | 489 | OD1 | ASP | A | 60 | 27.239 | 31.637 | -3.388 | 1.00 | 10.56 |
| ATOM | 490 | OD2 | ASP | A | 60 | 27.926 | 32.207 | -1.407 | 1.00 | 11.89 |
| ATOM | 491 | N | LEU | A | 61 | 28.180 | 27.469 | -1.910 | 1.00 | 11.29 |
| ATOM | 492 | CA | LEU | A | 61 | 27.158 | 26.395 | -1.785 | 1.00 | 13.55 |
| ATOM | 493 | C | LEU | A | 61 | 27.140 | 25.792 | -0.377 | 1.00 | 12.39 |
| ATOM | 494 | O | LEU | A | 61 | 26.076 | 25.569 | 0.204 | 1.00 | 10.30 |
| ATOM | 495 | CB | LEU | A | 61 | 27.403 | 25.337 | -2.873 | 1.00 | 13.67 |
| ATOM | 496 | CG | LEU | A | 61 | 26.455 | 24.132 | -2.929 | 1.00 | 12.05 |
| ATOM | 497 | CD1 | LEU | A | 61 | 26.375 | 23.553 | -4.341 | 1.00 | 12.93 |
| ATOM | 498 | CD2 | LEU | A | 61 | 26.930 | 23.031 | -2.000 | 1.00 | 16.59 |
| ATOM | 499 | N | LYS | A | 62 | 28.333 | 25.566 | 0.144 | 1.00 | 10.86 |
| ATOM | 500 | CA | LYS | A | 62 | 28.491 | 25.175 | 1.544 | 1.00 | 13.89 |
| ATOM | 501 | C | LYS | A | 62 | 27.916 | 26.199 | 2.531 | 1.00 | 13.74 |
| ATOM | 502 | O | LYS | A | 62 | 27.209 | 25.822 | 3.464 | 1.00 | 12.28 |
| ATOM | 503 | CB | LYS | A | 62 | 30.002 | 25.061 | 1.776 | 1.00 | 12.89 |
| ATOM | 504 | CG | LYS | A | 62 | 30.365 | 24.572 | 3.153 | 1.00 | 13.43 |
| ATOM | 505 | CD | LYS | A | 62 | 31.694 | 23.828 | 3.063 | 1.00 | 12.14 |
| ATOM | 506 | CE | LYS | A | 62 | 32.046 | 23.190 | 4.390 | 1.00 | 15.51 |
| ATOM | 507 | NZ | LYS | A | 62 | 33.369 | 22.534 | 4.287 | 1.00 | 17.63 |
| ATOM | 508 | N | LYS | A | 63 | 28.139 | 27.478 | 2.257 | 1.00 | 13.23 |
| ATOM | 509 | CA | LYS | A | 63 | 27.570 | 28.584 | 3.080 | 1.00 | 12.05 |
| ATOM | 510 | C | LYS | A | 63 | 26.048 | 28.526 | 3.104 | 1.00 | 11.13 |
| ATOM | 511 | O | LYS | A | 63 | 25.438 | 28.658 | 4.164 | 1.00 | 12.64 |
| ATOM | 512 | CB | LYS | A | 63 | 27.932 | 29.946 | 2.489 | 1.00 | 11.27 |
| ATOM | 513 | CG | LYS | A | 63 | 28.277 | 30.918 | 3.588 | 1.00 | 13.99 |
| ATOM | 514 | CD | LYS | A | 63 | 28.534 | 32.349 | 3.116 | 1.00 | 14.19 |
| ATOM | 515 | CE | LYS | A | 63 | 29.792 | 32.484 | 2.276 | 1.00 | 16.43 |
| ATOM | 516 | NZ | LYS | A | 63 | 29.892 | 33.917 | 1.896 | 1.00 | 15.02 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ATOM | 517 | N | HIS | A | 64 | 25.479 | 28.309 | 1.940 | 1.00 | 11.61 |
| ATOM | 518 | CA | HIS | A | 64 | 24.016 | 28.237 | 1.840 | 1.00 | 10.72 |
| ATOM | 519 | C | HIS | A | 64 | 23.445 | 26.952 | 2.461 | 1.00 | 10.86 |
| ATOM | 520 | O | HIS | A | 64 | 22.427 | 27.022 | 3.140 | 1.00 | 11.94 |
| ATOM | 521 | CB | HIS | A | 64 | 23.561 | 28.299 | 0.392 | 1.00 | 11.17 |
| ATOM | 522 | CG | HIS | A | 64 | 22.055 | 28.401 | 0.271 | 1.00 | 11.82 |
| ATOM | 523 | ND1 | HIS | A | 64 | 21.332 | 29.442 | 0.778 | 1.00 | 12.13 |
| ATOM | 524 | CD2 | HIS | A | 64 | 21.170 | 27.481 | -0.259 | 1.00 | 16.49 |
| ATOM | 525 | CE1 | HIS | A | 64 | 20.008 | 29.178 | 0.545 | 1.00 | 11.48 |
| ATOM | 526 | NE2 | HIS | A | 64 | 19.911 | 27.978 | -0.083 | 1.00 | 12.54 |
| ATOM | 527 | N | GLY | A | 65 | 24.189 | 25.846 | 2.403 | 1.00 | 9.84 |
| ATOM | 528 | CA | GLY | A | 65 | 23.765 | 24.640 | 3.124 | 1.00 | 10.13 |
| ATOM | 529 | C | GLY | A | 65 | 23.720 | 24.833 | 4.640 | 1.00 | 12.09 |
| ATOM | 530 | O | GLY | A | 65 | 22.785 | 24.356 | 5.310 | 1.00 | 10.90 |
| ATOM | 531 | N | VAL | A | 66 | 24.653 | 25.629 | 5.138 | 1.00 | 11.12 |
| ATOM | 532 | CA | VAL | A | 66 | 24.554 | 26.025 | 6.578 | 1.00 | 11.96 |
| ATOM | 533 | C | VAL | A | 66 | 23.297 | 26.851 | 6.899 | 1.00 | 10.82 |
| ATOM | 534 | O | VAL | A | 66 | 22.631 | 26.593 | 7.910 | 1.00 | 11.04 |
| ATOM | 535 | CB | VAL | A | 66 | 25.799 | 26.777 | 7.027 | 1.00 | 12.19 |
| ATOM | 536 | CG1 | VAL | A | 66 | 25.709 | 27.227 | 8.478 | 1.00 | 11.49 |
| ATOM | 537 | CG2 | VAL | A | 66 | 27.062 | 25.931 | 6.850 | 1.00 | 13.42 |
| ATOM | 538 | N | THR | A | 67 | 22.967 | 27.768 | 5.995 | 1.00 | 11.19 |
| ATOM | 539 | CA | THR | A | 67 | 21.780 | 28.628 | 6.142 | 1.00 | 11.60 |
| ATOM | 540 | C | THR | A | 67 | 20.497 | 27.825 | 6.258 | 1.00 | 9.84 |
| ATOM | 541 | O | THR | A | 67 | 19.740 | 28.004 | 7.218 | 1.00 | 10.87 |
| ATOM | 542 | CB | THR | A | 67 | 21.678 | 29.575 | 4.949 | 1.00 | 11.28 |
| ATOM | 543 | OG1 | THR | A | 67 | 22.640 | 30.627 | 5.157 | 1.00 | 12.09 |
| ATOM | 544 | CG2 | THR | A | 67 | 20.315 | 30.245 | 4.853 | 1.00 | 12.21 |
| ATOM | 545 | N | VAL | A | 68 | 20.397 | 26.839 | 5.373 | 1.00 | 11.34 |
| ATOM | 546 | CA | VAL | A | 68 | 19.193 | 25.996 | 5.309 | 1.00 | 9.46 |
| ATOM | 547 | C | VAL | A | 68 | 19.020 | 25.073 | 6.515 | 1.00 | 8.75 |
| ATOM | 548 | O | VAL | A | 68 | 17.966 | 25.083 | 7.152 | 1.00 | 10.55 |
| ATOM | 549 | CB | VAL | A | 68 | 19.215 | 25.219 | 3.994 | 1.00 | 9.32 |
| ATOM | 550 | CG1 | VAL | A | 68 | 18.113 | 24.174 | 3.924 | 1.00 | 12.56 |
| ATOM | 551 | CG2 | VAL | A | 68 | 19.100 | 26.200 | 2.837 | 1.00 | 10.55 |
| ATOM | 552 | N | LEU | A | 69 | 20.079 | 24.401 | 6.906 | 1.00 | 11.18 |
| ATOM | 553 | CA | LEU | A | 69 | 19.977 | 23.500 | 8.092 | 1.00 | 12.51 |
| ATOM | 554 | C | LEU | A | 69 | 19.814 | 24.216 | 9.418 | 1.00 | 10.58 |
| ATOM | 555 | O | LEU | A | 69 | 19.230 | 23.676 | 10.347 | 1.00 | 10.91 |

# DSSP: Dictionary of Secondary Structures in Proteins

**Sequence and secondary structure for 4MBN chain A**

```
1                VLSEGEWQLV LHVWAKVEAD VAGHGQDILI RLFKSHPETL EKFDRFKHLK
                    HHHHHHH HHHHHHHGGG HHHHHHHHHH HHHHH HHHH HT GGGTT

51               TEAEMKASED LKKHGVTVLT ALGAILKKKG HHEAELKPLA QSHATKHKIP
                 SHHHHHH HH HHHHHHHHHH HHHHHHTTTT   HHHHHHHH HHHHHTS

101              IKYLEFISEA IIHVLHSRHP GDFGADAQGA MNKALELFRK DIAAKYKELG
                 HHHHHHHHHH HHHHHHHH G GGS HHHHHH HHHHHHHHHH HHHHHHHHHT

151              YQG
```

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | A | G | H | > S+ | 0 | 0 | 35 | 2,-0.2 | 4,-1.6 | 1,-0.2 | -1,-0.2 | 0.823 | 107.4 | 48.1 | -63. | | 4.5 |
| 6 | A | E | H | > S+ | 0 | 0 | 51 | 2,-0.2 | 4,-1.8 | 1,-0.2 | -1,-0.2 | 0.883 | 109.7 | 52.9 | -77.1 | -34.6 | |
| 7 | A | W | H | X S+ | 0 | 0 | 15 | -4,-2.7 | 4,-2.6 | 2,-0.2 | 5,-0.3 | 0.894 | 105.3 | 56.2 | -63.6 | -34.4 | |
| 8 | A | Q | H | X S+ | 0 | 0 | 133 | -4,-2.1 | 4,-2.5 | 1,-0.2 | 5,-0.2 | 0.938 | 107.3 | 48.2 | -56.7 | -47.2 | |
| 9 | A | L | H | X S+ | 0 | 0 | 55 | -4,-1.6 | 4,-1.5 | 1,-0.2 | -1,-0.2 | 0.855 | 112.8 | 50.0 | -60.7 | -40.5 | |
| 10 | A | V | H | X S+ | 0 | 0 | 0 | -4,-1.8 | 4,-2.0 | 2,-0.2 | -1,-0.2 | 0.917 | 114.6 | 40.0 | -65.7 | -50.1 | |
| 11 | A | L | H | X S+ | 0 | 0 | 44 | -4,-2.6 | 4,-2.2 | 2,-0.2 | -2,-0.2 | 0.842 | 107.7 | 61.6 | -78.6 | -27.0 | |
| 12 | A | H | H | X S+ | 0 | 0 | 120 | -4,-2.5 | 4,-0.6 | -5,-0.3 | -1,-0.2 | 0.965 | 109.4 | 43.1 | -64.9 | -40.0 | |

# Propensity

**The propensity of an amino acid residue i in any conformation (helix or strand or turn or coil) has been defined as the percentage of residue i in that conformation to the percentage of all residues in the same conformation.**

**propensity$_\alpha$(i) = % of residue i in α-helix/ % of all residues in α-helix.**

**% of residue i in α-helix = $n_\alpha(i)/N(i)$**

**$n_\alpha$(i) = number of residues of type i in α-helix**

**N(i) = number of residues of type i in the whole dataset**

**% of all residues in α-helix = $n_\alpha/N$**

**$n_a$ = total number of residues in a-helix**

**N = total number of residues in the whole dataset**

# Propensity

```
VLSEGEWQLV LHVWAKVEAD VAGHGQDILI RLFKSHPETL EKFDRFKHLK
    HHHHHHH HHHHHHHGGG HHHHHHHHHH HHHHH HHHH HT GGGTT

TEAEMKASED LKKHGVTVLT ALGAILKKKG HHEAELKPLA QSHATKHKIP
SHHHHHH HH HHHHHHHHHH HHHHHHTTTT   HHHHHHHH HHHHHTS

IKYLEFISEA IIHVLHSRHP GDFGADAQGA MNKALELFRK DIAAKYKELG
HHHHHHHHHH HHHHHHHH G GGS HHHHHH HHHHHHHHHH HHHHHHHHHT

YQG
```

**E.g. Ala: % of Ala in α-helix = $N_\alpha$(Ala)/N(Ala) = 15/16 = 0.94**

**% of all residues in α-helix = $N_\alpha$/N = 115/153 = 0.75**

**Propensity of Ala = 0.94/0.75 = 1.25**

**Propensity of Gly:**

# Algorithm

1. Compute the occurrence of 20 residues in helix

2. Compute the occurrence of 20 residues in whole protein

3. Compute the ratio

4. Compute total number of residues in helix

5. Divide with total number of residues in a protein

6. Divide 3 by 5, to get the propensity of all the 20 amino acid residues in helix.

**Table 5.2** Chou–Fasman parameters

| Residue | $P_\alpha$ | Residue | $P_\beta$ | Residue | $P_t$ |
|---|---|---|---|---|---|
| Glu | Hα 1.53 | Hβ Met | 1.67 | Asn | 1.68 |
| Ala | 1.45 | Val | 1.65 | Gly | 1.68 |
| Leu | 1.34 | Ile | 1.60 | Ser | 1.56 |
| His | hα 1.24 | hβ Cys | 1.30 | Pro | 1.54 |
| Met | 1.20 | Tyr | 1.29 | Asp | 1.26 |
| Gln | 1.17 | Phe | 1.28 | Tyr | 1.25 |
| Trp | 1.14 | Gln | 1.23 | Cys | 1.17 |
| Val | 1.14 | Leu | 1.22 | Trp | 1.11 |
| Phe | 1.12 | Thr | 1.20 | Lys | 1.01 |
| Lys | Iα 1.07 | Trp | 1.19 | Arg | 1.00 |
| Ile | 1.00 | Iβ Ala | 0.97 | Thr | 1.00 |
| Asp | iα 0.98 | iβ Arg | 0.90 | Phe | 0.71 |
| Thr | 0.82 | Gly | 0.81 | His | 0.69 |
| Ser | 0.79 | Asp | 0.80 | Met | 0.67 |
| Arg | 0.79 | bβ Lys | 0.74 | Ile | 0.58 |
| Cys | 0.77 | Ser | 0.72 | Ala | 0.57 |
| Asn | bα 0.73 | His | 0.71 | Gln | 0.56 |
| Tyr | 0.61 | Asn | 0.65 | Leu | 0.53 |
| Pro | Bα 0.59 | Pro | 0.62 | Glu | 0.44 |
| Gly | 0.53 | Bβ Glu | 0.26 | Val | 0.30 |

**Hα(β): Strong helix (strand) former**
**hα(β): Helix (strand) former**
**Iα(β): Weak helix (strand) former**
**iα(β): Weak helix (strand) breaker**
**bα(β): Helix (strand) breaker**
**Bα(β): Strong helix (strand) breaker**

# Rules for identifying helix

**The propensity values and the residues belonging to these 12 classes are shown in Table 5.2.**

**For a protein sequence, assign the appropriate parameters from Table 5.2.**

**Helix:**

**the values of the six parameters are Ha = ha =1; Ia = ia = 0; Ba = ba = -1;**

**scan for window of 6 residues, where score ≥ 4, i.e. at least four helix formers and not more than one helix breaker;**

**extend the length in both directions until 4-residue window has the average propensity < 1;**

**continue the search and locate all helical regions in the sequence.**

**Sequence and secondary structure for 4LYZ chain A**

```
1                KVFGRCELAA AMKRHGLDNY RGYSLGNWVC AAKFESNFNT QATNRNTDGS
                  B  HHHHHH HHHHTT TTB TTB HHHHHH HHHHHHTTBS S EEE SSS

51               TDYGILQINS RWWCNDGRTP GSRNLCNIPC SALLSSDITA SVNCAKKIVS
                 EEETTTTEET TTT B SS T T   SS SBG GGGGSS  HH HHHHHHHHTT

101              DGNGMNAWVA WRNRCKGTDV QAWIRGCRL
                 TSSGGGGSHH HHHHTTTS G GGGSTT
```

**KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGS**
**TDYGILQINSRWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVS**
**DGNGMNAWVAWRNRCKGTDVQAWIRGCRL**

KVFGRC
VFGRCE
FGRCEL
GRCELA
RCELAA

# Rules for identifying β-strand

**The values of the six parameters are Hb = hb =1; Ib = ib = 0; Bb = bb = -1;**

**scan for window of 5 residues, where score > 3, i.e. at least three strand formers and not more than one strand breaker;**

**extend the length in both directions until 4-residue window has the average propensity < 1;**

**continue the search and locate all strand regions in the sequence.**

**Conflict situation:**

**a region containing overlapping helical and strand assignments is considered as a helix (or strand) if average propensity of a-helix (b-strand) is greater than that of b-strand (a-helix).**

**Choose:**

(A) Program, (B) Protein (sequence/accession) (C) Analyze protein:

(A) Program: Chou-Fasman Secondary Structure prediction

(B) Protein sequence: FASTA format

```
>4LYZ:A|PDBID|CHAIN|SEQUENCE
KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQIN
SRWWCNDGRTPGSRNLCNIPC
SALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKGTDVQAWIRGCRL
```

(C) Do Analysis — Submit Sequence

```
 CHOFAS predicts protein secondary structure
 version 2.0u66 September 1998
Please cite:
 Chou and Fasman (1974) Biochem., 13:222-245
```

---

```
 Chou-Fasman plot of @, 129 aa;
4LYZ:A|PDBID|CHAIN|SEQUENCE


              .         .         .         .         .         .
      KVFGRCELAAAMKRHGLDNYRGYSLGNWVCAAKFESNFNTQATNRNTDGSTDYGILQINS
 helix     <-------->           <-------->                   <>
 sheet                      EEEEEEE         EEEEEEE       EEEEEEEE
 turns T                                  T   T    T   T T        T

              .         .         .         .         .         .
      RWWCNDGRTPGSRNLCNIPCSALLSSDITASVNCAKKIVSDGNGMNAWVAWRNRCKGTDV
 helix                        <---------------->    <---->        <>
 sheet               EEEE                                        EEE
 turns      TT   T T            T         T    T T         T

      QAWIRGCRL
 helix
 sheet EEEE
 turns

 Residue totals: H: 47   E: 33   T: 16
        percent: H: 36.4 E: 25.6 T: 12.4
```

**http://fasta.bioch.virginia.edu/fasta/chofas.htm**